IntechOpen

Public Health

Social and Behavioral Health

Edited by Jay Maddock

PUBLIC HEALTH – SOCIAL AND BEHAVIORAL HEALTH

Edited by **Jay Maddock**

Public Health - Social and Behavioral Health
http://dx.doi.org/10.5772/2242
Edited by Jay Maddock

Contributors

Jay Maddock, David O'Riordan, Michelle Schneider, Claudio Nigg, Judith Väth, Syuichi Ooki, Kiyomi HIko, Adedapo Afolabi, Terence John Ryan, Stephen Ersser, Claire Fuller, Julian Alston, Abigail Okrent, Joanna C Parks, Teresa Cawley, Matthew Smith, Basia Belza, Mary Altpeter, Justin Dickerson, SangNam Ahn, Marcia Ory, Adam Davis, Mindy Benson, Jürgen Noeske, Steven Provyn, Nerilee Hing, Helen Breen, Thozama Mandisa Lutya, Mark Lanier, Kersti Pärna, Mariliis Tael, Inge Ringmets, Katrin Aasvee, Alberto Firenze, Steven Garasky, Craig Gundersen, Susan Stewart, Joey Eisenmann, Brenda Lohman, Amy Thompson, Joel Lamounier, Flávio Diniz Capanema, Daniela Rocha, Juyoung Lee, Qing Li, Liisa Tyrväinen, Yuko Tsunetsugu, Bum-Jin Park, Takahide Kagawa, Yoshifumi Miyazaki, Aida Turrini, Claudia Arganini, Raffaella Comitato, Fabio Virgili, Anna Saba, Ralphenia Pace, Melissa Johnson, Norma Dawkins, Nancy Chin, Scott Smith, Emine Kale, Bernadette Nirmal Kumar, Sandy Lazarus, Bonnie Duran, Leon Caldwell, Samed Bulbulia, Stefan Storcksdieck genannt Bonsmann, Stefaan De Henauw, Vera Verbestel, Ilse De Bourdeaudhuij, Laura Fernández Celemín, Staffan Marild, Katharina Gallois, Holger Hassel, Lucia Tramonte, Jon Douglas Willms

Notice

First published in Croatia, 2012 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019.
IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

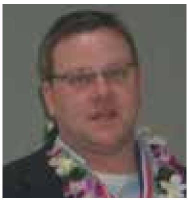
Public Health - Social and Behavioral Health
Edited by Jay Maddock

p. cm.

ISBN 978-953-51-0620-3

eBook (PDF) ISBN 978-953-51-4295-9

Meet the editor



Jay Maddock, Ph.D., FAAHB is Professor and Director of the Office of Public Health Studies at the University of Hawaiʻi at Mānoa and Luojia Chair Professor at Wuhan University. Professor Maddock has extensive experience in system, environmental and policy research to improve population level risk factors for chronic disease. Dr. Maddock chaired the state board of health and was a charter member of the NIH study section on Community-Level Health Promotion. He is an author of over 75 scientific articles. He is the Honorary Secretary for the Asia-Pacific Academic Consortium for Public Health. His research has been featured in several national magazines including Eating Well, Prevention and Good Housekeeping.

Contents

Preface

Human health is greatly influenced by the daily behaviors and patterns that make up our lives. Tobacco use, physical inactivity, poor nutrition, immoderate alcohol use, drug use, violence, unsafe sexual practices and other risky behaviors account for a large proportion of premature morbidity and mortality worldwide. Over the last couple of decades, the role of the social sciences in influencing and changing human behaviors has become more prominent. Psychology, sociology, political science, economics, anthropology, communications and political science have all played an important role in health counseling, group based interventions, social marketing and policy change. A student being trained in a Master's of Public Health program in Health Promotion needs to be versed in all of these areas to be effective at changes population level behaviors.

This book provides an overview of the influence of the social and behavioral sciences to many public health issues that confront us today. In the first section, the chapters explore the growing problem of obesity and the related behavioral factors of physical inactivity and poor nutrition. Chapters examine the effects of food policies including iron fortification of foods, psychological theory testing to improve health, gender differences, the effect of stress on obesity and strategies to prevent childhood obesity and reach rural communities. In the second section, the chapters explore the effects of addictive behaviors. Issues around alcohol use, drugs and gambling are explored both in comprehensive reviews and in county level analyses. The third section examines a variety of different approaches and methods to changing health behaviors. These include evidence-based public health, community-based participatory research, empowerment, communication and dialog and even nature therapy. The final section reviews a variety of at-risk populations including prisoners, men who have sex with men, school children, deaf persons, school children and young women involuntarily participating in prostitution. Reviews of important but often neglected public health areas such as gun violence, skin care for all and asthma are also presented.

This book exemplifies the global nature of public health. All six inhabited continents are represented by authors in this book. The home country of the authors include Australia, Estonia, South Africa, Nigeria, Brazil, Canada, Korea, Finland, Japan, Great Britain, Ireland, USA, Belgium, Sweden and Italy. This trans-national list of authors provides an important view of the future of public health and the increased need to collaborate with public health professionals across the world to address the myriad of public health issues. I hope you enjoy reading the following chapters. I find them to be insightful and to provide an excellent collection of the ways that the social and behavioral health sciences are being used to protect and promote the public’s health. Aloha.

Prof. Jay Maddock
Department of Public Health Sciences,
University of Hawai‘i at Mānoa
USA

Section 1

Obesity, Food and Physical Activity

1

The IDEFICS Intervention Toolbox – A Guide to Successful Obesity Prevention at Community Level

Vera Verbestel[1], Stefaan De Henauw[1], Staffan Marild[2],
Stefan Storcksdieck genannt Bonsmann[3], Laura Fernández Celemín[3],
Katharina Gallois[4], Holger Hassel[4,5] and Ilse De Bourdeaudhuij[1]
[1]*Ghent University, Ghent,*
[2]*Department of Public Health and Community Medicine, Gothenburg,*
[3]*European Food Information Council (EUFIC), Brussels,*
[4]*University of Bremen, Bremen,*
[5]*Hochschule Coburg, University of Applied Sciences, Coburg,*
[1,3]*Belgium*
[2]*Sweden*
[4,5]*Germany*

1. Introduction

The chapter provides an overview of the IDEFICS (Identification and prevention of Dietary- and lifestyle-induced health EFfects In Children and infantS) intervention and its general content and structure, including the core set of intervention modules, communication strategies and corresponding standard operating procedures for use in preschools, primary schools and other settings and dissemination channels.

The chapter does not present information on the development of the IDEFICS intervention but aims to provide practical guidance and recommendations for local policy makers and/or local public health authorities who wish to implement the intervention in their cities or communities. Every authority or institution wishing to implement the intervention will have to adapt all the intervention modules to account for local and/or culture-specific constraints. This can be done on the basis of qualitative research or equally valid sources of relevant information. General aspects of the IDEFICS project and the development of the IDEFICS community-oriented intervention programme have been described in detail elsewhere (Ahrens et al., 2006; Ahrens et al., 2011; De Henauw et al., 2011; Pigeot et al., 2010; Verbestel et al., 2011).

1.1 Timing of the IDEFICS intervention activities

The timeline of the IDEFICS intervention is divided into three parts, referring to the classical phasing of establishing interventions (see Table 1): intervention adoption phase, intervention implementation phase and intervention dissemination phase. The intervention adoption phase was proceeded by a one-year preparation phase.

Intervention preparation phase (Year 1)

The implementation of a culturally adapted version of the IDEFICS intervention needs to be planned in advance. Therefore, local authorities or institutions aiming to implement the intervention, are recommended to consider a preparation phase of at least one year. During this period, necessary arrangements can be made for the start of the adoption phase.

One of the most important arrangements that needs to be made during the preparation phase is to build a local intervention team. A local intervention team is a group of people that is preferably composed by the local authority or institution that aims to implement the intervention. It should consist of local experts in the field of health promotion and/or representatives of the respective authority or institution. The local intervention team needs to be able to prepare the adoption phase of the intervention and to support and supervise the implementation of the intervention during the first year. This support and supervision will gradually decrease so that local structures (community, family and schools) can independently continue the health-promoting efforts initiated by the intervention. The role of the local intervention team throughout the phases of the intervention is described in more detail below.

First intervention period = Adoption phase (Year 2)

Assuming that a school year starts in September, this phase should cover the period from September to August of Year 2 and starts right after the preparation phase (Year 1). During this period, the intervention will be launched and installed in the community, family and schools. This period is characterised by the continuous provision of material and logistic input, support and supervision from the local intervention team.

	Year 1	Year 2	Year 3	Year 4 and onwards
Sep	Preparation phase	Intervention adoption phase (support and supervision)	Intervention implementation phase (only supervision)	Intervention dissemination phase (no support, no supervision)
Oct				
Nov				
Dec				
Jan				
Feb				
Mar				
Apr				
May				
Jun				
Jul				
Aug				

Table 1. Timeline of the IDEFICS intervention activities

Second intervention period = Implementation phase (Year 3)

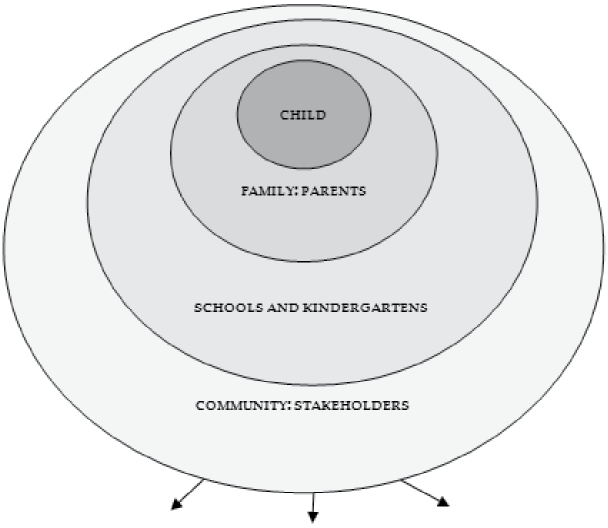
This phase covers the period from September to August of Year 3 and starts right after the adoption phase (Year 2). During this period, the intervention will be sustained and further progressed by the community, schools and families themselves without continuous material support from the local intervention team, but still with some degree of supervision.

Third intervention period = Dissemination phase (Year 4 and onwards)

This period starts in September of the school year after the implementation phase (Year 3) and aims at continuing the IDEFICS intervention without support and supervision of the local intervention team.

1.2 Dimensions of the IDEFICS intervention

The development of the child has to be viewed from an interactive and contextual perspective. The ecological environment of a child includes the family and the school which are in turn situated in the community and the society at large. Interactions within and among these social contexts result in changes within, and influence the development of, the individual child (Davison & Birch, 2001).



- Recreation and fitness sector
- Sport and youth organisations
- Private actors (food companies, grocery stores)

- Residents' associations
- Organisations targeting persons with low socio-economic status
- Social services and welfare sector
- Communication sector: local media

- Local municipality (especially public health authorities)
- Local politicians
- Health care providers (paediatrician, family doctor, …)

Fig. 1. Dimensions of the IDEFICS intervention: the individual, the family, the school and kindergarten and the community level

The community-based IDEFICS intervention takes a holistic approach associated with this contextual and interactive perspective of human development and is being executed on three intertwining levels: community level, school level and family level (Figure 1). Possible stakeholders in the community that can have an impact on the prevention of obesity are illustrated. The local intervention team has to analyse its own intervention community and identify the key persons and stakeholders.

1.3 Behaviours targeted by the IDEFICS intervention

The behaviours that were targeted by the IDEFICS intervention can be found in Table 2. From this point onwards, these target behaviours will be called the "key messages".

The selection of these key messages was based on the available evidence in the scientific literature. Detailed information on this selection process is outside the scope of this chapter and can be found elsewhere (Verbestel et al., 2011). Due to the complex nature of the problem, there is also scientific evidence available showing additional behaviours having an influence on the development of childhood obesity (e.g. portion sizes and snacking). This means that multi-topic interventions for the prevention of childhood obesity do not necessarily have to focus on the below mentioned key messages. Other behaviours can be chosen as the focus of the intervention, as long as they are supported by scientific evidence in the childhood obesity preventive literature.

Diet	Physical activity	Stress, coping and relaxation
Stimulate daily consumption of water	Reduce TV-viewing	Spend more time together
Stimulate daily consumption of fruit and vegetables	Stimulate daily physical activity	Ensure adequate sleep duration

Table 2. The six key messages targeted by the IDEFICS intervention

1.4 Overview of the IDEFICS intervention modules

The identification and elaboration of the IDEFICS intervention modules was predominantly carried out during the first year of the project on the basis of literature reviews, expert consultations and focus group research (Verbestel et al., 2011). The modules have been developed as distinct sets of activities but it should be realised that some of these modules are partly overlapping and/or interacting with one another. Table 3 illustrates how different modules have been conceived within a grid of targeted behaviours and intervention levels. Some modules have a more general scope whereas others are much more specific and focusing on a particular intervention level or behavioural dimension. The general content and structure supplemented with a description of the core actions for all the IDEFICS intervention modules is provided in the sections below.

	COMMUNITY	SCHOOL	FAMILY	INDIVIDUAL
DIET	Module 1 Module 2 Module 3	Module 4 Module 8 Module 9	Module 10	Module 5
PHYSICAL ACTIVITY	Module 1 Module 2 Module 3	Module 4 Module 6 Module 7	Module 10	Module 5
STRESS, COPING AND RELAXATION	Module 1 Module 2 Module 3	Module 4	Module 10	Module 5

Module 1: Involvement of community partners
Module 2: Long-term media campaign and public relations strategy
Module 3: Lobbying for community environmental and policy interventions
Module 4: Building partnerships
Module 5: Education of children
Module 6: Environmental changes related to physical activity – The Active Playground
Module 7: Health-related physical education curricula
Module 8: Environmental changes and school policy related to water consumption
Module 9: Environmental changes and school policy related to fruit and vegetables
Module 10: Education of parents

Table 3. Overview of the IDEFICS modules by behavioural focus and intervention level

2. IDEFICS intervention modules at community level

2.1 Module 1: Involvement of community partners

All the community partners will be engaged in the intervention by means of a community platform, i.e. a working group on meta (community) level in which all relevant stakeholders need to be represented. The local intervention team is an essential leading actor in the implementation of this module as it is responsible for triggering and coordinating the establishment and the operation of the community platform. This means that the implementation quality of this intervention module strongly depends on the leadership capacity of the local intervention team.

Objectives

- Create involvement and commitment of all relevant sectors of the community.
- Make it possible to implement the intervention at community level by the combination of support from the intervention team in the early phase followed by a gradually increasing independence of the stakeholders in the community.
- Stimulate the community to develop, organise and promote programmes and structural changes that encourage the healthy behaviours targeted in the intervention.
- Prepare the dissemination phase of the community-based intervention.

Establishment and composition of the community platform

- The community platform has to be established, organised and coordinated by a local intervention team.
- Preferably, one community platform is created in the community. When it is obvious that the creation of a single community platform is not feasible, it can be envisaged to create more than one community platform (e.g. when different communities are part of a larger intervention region).
- If a community already has a community platform, it is recommended to integrate the IDEFICS platform within the existing one.
- Every community platform has to include at least one representative from all important stakeholders in the community: each local intervention team has to analyse their own community to identify such key persons within these stakeholders. Stakeholders are perceived important when they are able to reach vulnerable groups, persons with low(er) socio-economic status or migration groups and/or if they can contribute significantly to the prevention of (childhood) obesity.

Task and responsibilities of the community platform

The community platform is responsible for the development and implementation of all the intervention modules at community level, i.e.:

- Module 2: Long-term media campaign and public relations strategy
- Module 3: Lobbying for community environmental and policy interventions

Within this section, the tasks and responsibilities of the community platform are briefly outlined. More detailed and concrete descriptions can be found in the relevant module sections.

Because of the ecological perspective associated with the IDEFICS intervention, interactions within and among the different contexts in the community - as illustrated in Figure 1 above - are essential. The community platform is therefore expected to support the implementation of modules at other levels:

- Module 5: Education of children (school level)
- Module 10: Education of parents (family level)

The community platform will not be responsible for the intervention modules at the school level. The implementation of these modules will be organised by a working group at the school level (see module 4). Regarding module 5 and 10, the community platform will mainly provide logistic and reinforcing support (e.g. the provision of posters to the school working group). This support is explained in more detail within the appropriate module sections.

Operation of the community platform during the intervention adoption phase (Year 2)

The first part of the adoption phase (September – June, Year 2) is dedicated to building the infrastructure for implementing the intervention modules:

- All the participants of the community platform will receive instructions and guidelines about the modules that have to be implemented in the community.
- The community platform will work closely with the local intervention team in order to implement the community-level modules.

- The local intervention team will be in charge and will coordinate the community platform. If possible, the local intervention team is recommended to appoint a local coordinator from the beginning of the platform. The local coordinator is a person from the community (e.g. the chair of an already existing platform) who will be responsible for chairing the platform the year afterwards.
- The local intervention team and the community platform should preferably meet once a month. These meetings make it possible to evaluate the implementation of the modules and to discuss and solve practical problems that have occurred or possibly will occur.

Guidelines for the meetings of the community platform

- During the first meeting:
  - make an inventory of the local initiatives related to the prevention of obesity in the different sectors involved in the platform.
- During all the meetings:
  - always write a meeting report, as these reports can be used to analyse the implementation process of the intervention.
  - evaluate the aspects of the intervention modules that have been executed and provide strong support for the implementation of the (parts of the) modules that still have to be performed.
  - do not only discuss and evaluate the aspects of the intervention that have been successful but also address any challenges or failures.
  - continuously observe and detect what is going on inside the community related to the prevention of childhood obesity.
  - do not ignore but take notice of new initiatives that are proposed by the community platform members.

The second part of the adoption phase (July – August, year 2) is the transition period between the intervention adoption phase and the intervention implementation phase. This transition period is an intermediate stage between the intensively supported operation and the supervised operation of the community platform. During the transition period, following actions are recommended:

- In cooperation with the local intervention team, the community platform can search and appoint a local coordinator who can continue the responsibilities of the local intervention team. This local coordinator becomes the person in charge and will be responsible for coordinating the community platform.
- The local intervention team is responsible for the transfer of information to the local coordinator. It is essential that local coordinators can start their activities at the beginning of Year 3 (see Table 3).

Operation of the community platform during the intervention implementation phase (Year 3)

The community platform is expected to continue with the activities and to work out new initiatives with minimal supervision and without continuous support of the local intervention team.

From September of Year 3 onwards, the local intervention team no longer participates in the monthly meetings of the community platform. Starting from this moment, the local coordinator (of the community platform) has to be able to take over the role of the local intervention team.

The local intervention team must keep itself informed about the initiatives of the community platform. For this reason, the local intervention team and the community platform are recommended to meet 3 times during the intervention implementation phase, i.e. once between September and December of Year 3 and twice between January and August of Year 3.

Between the obligatory meetings, the community platform should have the possibility to keep in touch with the local intervention team in order to solve practical problems or get advice if necessary.

Operation of the community platform during the intervention dissemination phase (Year 4 onwards)

From year 4 onwards, the community platform is expected to operate completely independently, without any support or supervision from the local intervention team.

2.2 Module 2: Long-term media campaign and public relations strategy

A first topic in the long-term media campaign and public relations strategy is the overall approach by which the intervention will become well known and the key messages will be spread in the community. A second issue in module 2 is the specific promotion campaign for the key messages by means of a slogan intervention.

2.2.1 Multimedia and public relations campaign (overall strategy)

Objectives

- Facilitate cooperation of the stakeholders and community members with the IDEFICS intervention team.
- Avoid objection and resistance against the intervention.
- Inform all stakeholders and community members about the intervention.
- Attract funding or sponsoring.

The local intervention team can develop its own public relations strategy and timing, depending on the local needs and resources. Some examples of multimedia and public relation instruments that can be useful to fulfil the objectives of the overall strategy are leaflets (newsletters), information events, posters and a website in the local language. It is also recommended to not only rely on contacts and infrastructure of universities or health institutions but to establish own local media relations by developing for example a media kit (including press kit, contact list, media server) and/or organise media briefings (e.g. local kick-off event in September of Year 2).

2.2.2 Promotion campaign for the key messages (slogan intervention)

Objectives

- Make the community aware of the key messages.
- Promote the key messages as important components of long-term health.

Window and street posters as a promotion campaign for the key messages

Window and street posters can be used to promote the key messages within the community:

- Recommended characteristics of the window and street posters are described in Table 4 below.
- The following guidelines are recommended for the development of the window and street posters:
  - The local intervention team is recommended to develop 3 posters: 2 posters about physical activity and diet (one key message of each theme combined on each poster) and 1 poster about sleep duration.
  - The window and street posters should reflect the same message. The street poster should be a bigger copy of the window posters.
  - The posters should be simple and attractive: the message should be readable and understandable in a couple of seconds and the text should be a short, striking and attractive slogan.
  - Multi-colour printing is very expensive and may distract the attention of the message. Black and white printing is much cheaper and more clearly visible. A message in black and white printing (can be printed on coloured paper) has the reputation to be very effective.

	Window posters	**Street posters**
Dimensions of the poster	Minimum A3-format (maximum 50 - 70 cm, however, this size can already be a barrier to hang up the poster)	Minimum 50 – 70 cm
Target groups	Pedestrians, cyclists and car drivers	Pedestrians, cyclists and car drivers
Preferred places	Very suitable to hang up in public places, supermarkets and grocery stores, libraries, houses of residents, …	On the street site, preferably on places where a lot of people have to stop (e.g. traffic lights)

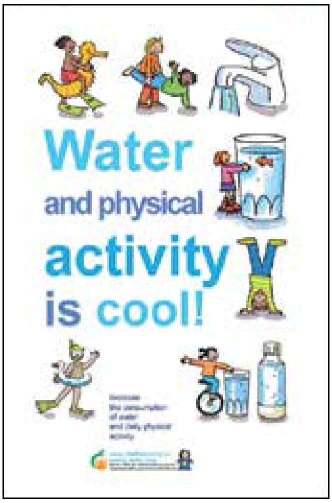
Table 4. Recommended characteristics of the window and street posters

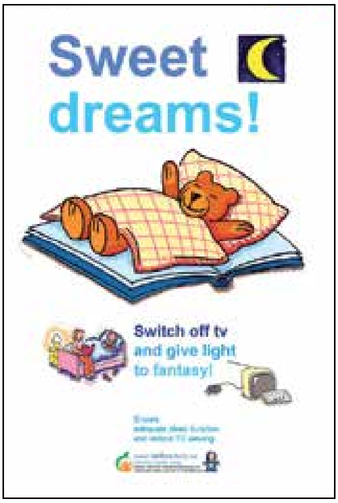
Because the visibility of the window and street posters will be synchronised with the integration of the key messages in the school curriculum (see module 5), the turnover of the posters is recommended to be set at 4 months for physical activity and diet and at 3 months for stress, coping and relaxation. The timing and turnover of the window posters during Year 2 of the IDEFICS project are shown in Table 5.

<table>
<tr><th colspan="11">Intervention adoption phase (Year 2)</th></tr>
<tr><td>Oct</td><td>Nov</td><td>Dec</td><td>Jan</td><td>Feb</td><td>Mar</td><td>Apr</td><td>May</td><td>Jun</td><td>Jul</td><td>Aug</td></tr>
<tr><td colspan="4">Window (and street) poster related to physical activity and diet
(1st poster)</td><td colspan="4">Window (and street) poster related to physical activity and diet
(2nd poster)</td><td colspan="3">Window (and street) poster related to sleep duration
(3rd poster)</td></tr>
</table>

Table 5. Timing and turnover of the window (and street posters) in Year 2 of the IDEFICS project

The window posters that were used in the IDEFICS intervention, are displayed below. These posters are available in the languages of those countries where the IDEFICS intervention was tested for its effectiveness (Belgium, Cyprus, Estonia, Germany, Hungary, Italy, Spain and Sweden).





Tasks and responsibilities of the community platform

- The distribution of the window and street posters is a task of the community platform. The community platform is responsible for the distribution of the posters at the right time in the different sectors of the community (see stakeholders). For example: the school is an important and easy setting to distribute the window posters to the residents of the community. Thus, the community platform also has to take care that the schools received the window posters in time.
- It is important that the community platform not only distributes the posters but also checks if the posters are in actual fact displayed. The community platform is responsible for encouraging the concerned institutions and/or persons to make the posters visible.
- The community platform is recommended to carry out a sample survey: this means that it has to register how many houses of the selected streets there is a poster in the window (the sample survey can be used for assessing the process of implementation of this module).

2.3 Module 3: Lobbying for community environmental and policy interventions

Module 3 requires the community platform to lobby for improving the community environment and for policy interventions to prevent obesity in the community. This task consists of four separated parts:

1. Community environmental interventions to promote physical activity
2. Community environmental interventions to promote the consumption of water
3. A short-term perspective of community-based prevention of obesity
4. A long-term perspective of community-based prevention of obesity

The community environmental interventions to promote physical activity and the consumption of water are strongly recommended to be put on the agenda of the community platform and should be implemented during the adoption phase.

The short-term perspective of community-based prevention of childhood obesity requires efforts of the community platform to undertake actions/activities that contribute to the prevention of obesity within the adoption phase.

The long-term perspective of the intervention aims at triggering new initiatives during the intervention adoption phase. Therefore, the community platform should start to advocate for environmental and policy interventions as soon as possible but the implementation of these interventions is not intended to be accomplished in the intervention adoption phase.

2.3.1 Community environmental interventions to promote physical activity: establishment of 'play streets' and community playgrounds

Objective

- Provide opportunities and possibilities for outdoor activity and outdoor play to the children in the community areas at risk (= areas without opportunities/possibilities to play outdoors).

Concept of play streets

Because the time children spend outdoors is positively correlated with higher physical activity levels in children (Ferreira et al., 2007), it is important that all children in the community at least have the possibility to be active outside. Community playgrounds are very important and the most favourable way to promote playing outside among young children. However, the establishment of community playgrounds is not possible everywhere. In such cases, play streets can provide children with safe opportunities to be active outside the home and attract children to different recreational activities. Play streets are spaces within neighbourhoods where road space is made available for children's play on weekend days and/or holidays (streets are closed for traffic during that time).

The organisation of play streets is mostly focused on specific areas in the community with few options for leisure activities and/or opportunities to spend active time outdoor. It is supposed that families with a low(er) socio-economic status live in neighbourhoods with busy streets and no gardens, thus lack safe structures to play outside. Play streets are an easy way to remedy this situation.

Identification of the concept of play street in the community

The local intervention team has to identify if the concept of play streets already exists in the community. In some countries, the concept of play streets is already well known.

- If the concept DOES exist in the community, the standard operating procedure should be used.
- If the concept DOES NOT exist in the community, the local intervention team and the community platform are responsible for launching the play streets in the community in strong cooperation with the local municipalities.

Identification of the community areas "at risk"

The community platform has to identify the areas in the community where children have no or not enough opportunities to spend time outdoors. This can be done by identifying the number of community playgrounds, their location and condition (in areas where children have enough opportunities for outdoor activity) and by identifying the areas in the community where play streets need to be established (in areas with no or not enough places for outdoor activity).

Guidelines for the establishment of the play streets or safe playgrounds in the selected areas

- Inform residents in the selected streets about the concept of play streets.
- Try to convince them to organise a play street in the street that they live.
- Motivate one resident to be the person in charge for a specific play street: one resident of the street has to act as an intermediary between the community platform and the residents of the street.
- If it is not possible to establish a play street in a certain area, try to use a public place (e.g. parking grounds) as a play street and promote this initiative in the neighbourhood.
- Play streets could be organised on weekends (Sundays) and particularly during vacation periods and holidays, throughout the intervention period and beyond.
- Advocate for the restoration of existing community playgrounds to a reasonable condition. For regions without any possibility for outdoor activity, advocate for new community playgrounds.

2.3.2 Community environmental interventions to promote the consumption of water: installation of water fountains in public places

Objective

- Provide the residents of the community with opportunities to drink water in public places.

Tasks and responsibilities of the community platform

- The community platform is responsible for the availability and the promotion of water fountains in public places in the community.
- The community platform has to advocate for the placement of water fountains/dispensers in public places (e.g. public library, sports centres, squares).
- The community platform has to advertise the water fountains/dispenser in the community so that people are encouraged to use them (e.g. promote it on the website of the community, in the school and/or community paper, in the local newspapers).

2.3.3 Short-term perspective of community-based prevention of childhood obesity

Every stakeholder represented in the community platform should try to undertake activities related to the prevention of obesity during the intervention adoption phase. Table 6 is a non-exhaustive list of possibilities that can be pursued by the stakeholders of the community. It is recommended that the stakeholders make efforts to realise some of these or similar initiatives.

Stakeholder	Possible actions
Local municipality (public health authorities) and local politicians	- Provide water and fruit and/or vegetables during internal meetings (e.g. meetings of the town council) and public meetings (e.g. assembly of political parties). - Stimulate the employees of the local authorities to drink water and provide free water during working hours. - Stimulate the employees of the local authorities to eat fruit and/or vegetables as a snack during working hours.
Private sector (food companies, grocery stores)	- Organise shopping tours, grocery tastings, cooking demonstrations, nutrition labelling. - Promote water and healthy food products such as fruit and vegetables. - Provide easy recipes with fruit and/or vegetables that are typical for a certain season. - Provide ideas to drink water in several ways (e.g. with mint leaves, pieces of apple, …).
Working groups of the schools/kinder-gartens	- Organise extracurricular physical activity programmes. - Distribute information about sports and recreation programmes in the community. - Enable sports and recreation programmes to make use of the school facilities outside the school hours. - Organise active after-school programmes.
Sport and youth organisations	- Provide and promote free water during the activities. - Stimulate the children to not bring sugar-sweetened beverages. - Stimulate the children to bring fruit and/or vegetables instead of less healthy snacks. - Organise activities in which the family of the children can participate (family events).
Health care providers	- Provide assessment, counselling and referral on physical activity, diet, stress, coping and relaxation as part of health care. - Encourage parents to be role models for their children in the field of physical activity, diet, stress, coping and relaxation.

Table 6. List of possible actions that can be undertaken by the stakeholders of the community as part of the short-term perspective of community-based prevention of childhood obesity

2.3.4 Long-term perspective of community-based prevention of childhood obesity

In addition to the short-term perspective, the IDEFICS intervention also considers a long-term perspective in the prevention of childhood obesity. The start of the intervention is the best moment to start this process. The community platform should advocate for environmental and policy interventions that contribute to the prevention of childhood obesity. Based on the examples provided by Crawford & Jeffery (2005), Table 7 provides a list of possible initiatives that can be accomplished within the long term. These initiatives are linked to three different actors in the community which potentially have the ability to execute these proposals. The community platform should decide on the number of initiatives and trigger their execution by working with the relevant actors.

Stakeholder	Possible initiatives
Local municipality (public health authorities) and local politicians	- Identify national obesity prevention plans and provide a significant contribution. - Take initiatives that enable children to have access to sports and recreation programmes and the equipment and supplies that are needed to participate in such programmes. - Promote indoor activities instead of screen-based activities (e.g. ice-skating, indoor swimming, …). - Organise and promote programmes that stimulate walking, cycling, and the use of sports and recreation facilities in the community. - Promote local activities that provide options for (un)structured play for children in a safe environment and at minimal cost. - Enable the local public transport system to stop at the local swimming pool so that children can get there without any risk and additional cost. This applies also to other sports infrastructures. - Develop safe roads in the municipality, especially those leading to schools. Safe roads are those that have safe pavements, bicycle paths, trails, and crosswalks that facilitate walking and cycling. - Provide physical activity equipment into neighbourhoods that do not have access to physical activity and recreation facilities.
Private sector (food companies, grocery stores)	- Include healthy alternatives in the menus that are specifically available for children (e.g. include fruit as a dessert). - Provide and promote healthy foods (e.g. fruits and vegetables).
Working groups of the schools and kindergartens	- Make healthy foods available, accessible and attractive in the school environment. - Create price incentives or use cross-subsidies to facilitate and promote healthy food. - Remove sugar-sweetened beverages from vending machines in the school environment and replace them with water and/or other healthy options, or water dispensers. - Create a school nutrition policy that promotes a healthy diet. - Prohibit food advertising at school (e.g. sports sponsorships, exclusive marketing contracts to sell food and beverages) and do not start industry-sponsored collaborations. - Promote, enable and facilitate active commuting to schools (e.g. organise walk/bike to school days, organise walking school buses or bicycle trains, provide safe bicycle sheds). - Take care that adult and trained crossing guards are available at important and busy intersections around the school so that children can safely cross the streets on and from their way to school.

Table 7. Overview of possible initiatives for the long-term perspective of community-based prevention of childhood and adult obesity (Crawford & Jeffery, 2005)

3. IDEFICS intervention modules at school and kindergarten level

3.1 Module 4: Building partnerships

Building partnerships in the participating schools and kindergartens contains three levels of action:

1. Alerting the staff members of the schools and kindergartens to the intervention
2. Creation of a school working group
3. Creation of a school platform at community level

By means of the first part of module 4, the local intervention team will aim for teachers to take part in the intervention and to support the overall content of the project in the schools and kindergartens.

The second part of module 4, the creation of a school working group, is intended to create involvement and commitment among staff members and to facilitate the implementation of the intervention. Good cooperation with all staff members in all participating schools is the basic principle of this module.

The creation of a school platform makes it possible to gather all school working groups. This third part of module 4 creates a structure in which the schools can exchange knowledge and experiences, share a collective opinion and together elaborate on and start new initiatives to prevent childhood obesity.

3.1.1 Alerting staff members of the schools and kindergartens to the IDEFICS intervention

Objectives

- Inform the employees about the community-based intervention and particularly about the intervention that will take place at school level (aim, content, manual, guidelines).
- Increase awareness of the health behaviours which are advocated in the intervention.
- Motivate and encourage staff members to take part in the intervention and specifically in the school working group (explained in the second part of module 4).

Get in contact with the principals of the schools

- Inform them about the aim and content of the intervention.
- Receive an agreement for participation in the intervention.
- Make a first appointment with the schools which agree to participate or with schools which need some time for reflection and want more practical and concrete information.

Inform coordinating school organisations

If several different schools are gathered in a coordinating school organisation, it is suggested to inform this organisation that the local intervention team will contact each of their schools individually.

Make ample use of coordinating school organisations because they can play an important role in stimulating and supporting their schools during the implementation of the intervention. Additionally, they can be a starting point and negotiation partner in order to set up a school platform at community level.

Organise a first appointment with the principals

- Explain the intervention (receive an agreement for participation in the intervention if still necessary).
- Discuss possibilities to organise a meeting for staff members during which they will be informed about the intervention.
- You can use an introduction booklet as an incentive to stimulate the schools. An overview of the possible content of this booklet is provided in Table 8 below.

Documents	**Aim and content of the documents**
Brief overview of the intervention	Document which describes the aim and content of the project and which puts the project in a broader social context.
Manuals for each module prepared by the local intervention team	The manuals provide, for each module, the culturally adapted ideas which can be used by the school working groups in order to work out the intervention modules in their school.
PowerPoint Presentation	A presentation based on the document which briefly outlines the intervention. This presentation can be used by the school to inform parents or a third party about the intervention.
Article for the school newspaper	An example of an article that can be used to include in the school newspaper. This article informs the parents about the fact that the school participates in the intervention.

Table 8. Possible content of the introduction booklet that can be used during the first contacts with the schools

Organise a meeting for all staff members of a participating school

After having informed the school principals, it is important to alert the staff members of the schools or kindergartens. Therefore, the local coordinating team has to organise an information session for all employees of each school or kindergarten that will participate in the intervention. The following guidelines can be used for the organisation of this meeting:

- Organise the first meeting in April or May of Year 1 at the latest (assuming that the school year starts in September).
- Organise the meeting in strong cooperation with the school board to make clear that the school board supports the project.
- The information session for the staff members should be a promotion campaign for participation in the school working group and must trigger the successful formation of the group (detection of staff members that are interested in participating in the school working group).
- During the meeting, the staff members must receive explanations about the community-based intervention and the intervention activities that will take place at school level (aim, content, manual, guidelines).

3.1.2 Creation of a school working group

Objectives

- Involve staff members in the implementation of the school-based intervention modules.
- Implement the school intervention by combining a certain degree of support from the local intervention team and a certain independence from the school or kindergarten.

- Encourage the school or kindergarten to develop, organise and promote programmes that support the health behaviours advocated in the intervention.
- Lay the foundation of the dissemination phase of the intervention in the schools and kindergartens.

Establishment and composition of the school working group

The following strategy, guidelines and recommendations can be used to set up a school working group in every participating school/kindergarten.

General recommendations

- A school working group must be established in every participating school/ kindergarten.
- It is the responsibility of the local intervention team to initiate the establishment of the school working groups. When a school or kindergarten already has a school working group, try to cooperate with it.
- Motivation of the staff members should not be the sole criteria when selecting participants for the school working group. Potential impact of the staff member (position in the organisation, ability to implement actions advocated by the intervention, power) should also be considered.
- Every school working group should contain the following representatives:
  - Representative(s) of the school board (most important link)
  - Teachers whose field of study is related to the content of the intervention
  - Educator(s): persons who are responsible for supervision during free time and recess
  - Representative(s) of the parents' council
  - Representative(s) of staff members who have the ability to reach children at risk of developing excessive body fatness (especially those with obese parents and low(er) socio-economic status)
- Every school working group has to consist of at least 2 persons (local coordinator not included)

Inventory of motivated staff members who wish to participate in the school working groups

After the meeting, all staff members of the school receive an information letter describing the aim and content of the intervention and the responsibilities of a person engaging in the school working group. This letter should be formulated by the local intervention team in cooperation with the school board. The letter has to make clear that staff members can also engage in the school working group as a co-worker. Co-workers are contact persons for the school working group which can be asked for the organisation or support of certain activities related to the intervention. Staff members who are motivated to participate in the school working group or as a co-worker can present themselves to the principal of the school.

Appointment of a project leader in the school

The school working group is responsible for the implementation of the intervention. With the implementation and dissemination phase of the intervention in mind, it is advisable that someone, besides the local intervention team, takes the lead in the entire process. For this reason it is suggested to appoint a project leader in every participating school. It is up to the principal to decide who will be the project leader in the school.

- This person has to compose the school working group with staff members who wish to engage in implementing the IDEFICS intervention. It is not expected from the local intervention team to take the lead in the organisation of the school working groups in every participating school.
- The school project leader can be a person from an existing working group, another staff member or the principal him- or herself. It is important that this person is in close contact with the school, school board, staff members and children.
- The school project leader will be the contact person for the local intervention team.

Despite the fact that the school working group will be composed by the school project leader, the local intervention team is expected to stay in close contact with the school project leader in order to stimulate/support the organisation and composition of the school working group.

If the kindergarten and primary school are gathered in the same school and thus both belong to the same school board, it is recommended to set up two project leaders and two school working groups. This is necessary to be able to implement an intervention which is adapted to the different age groups that are part of the intervention.

Tasks and responsibilities of the school working group

The school working group will be responsible for the organisation of a cooking and activity competition in their own kindergarten or school and the implementation of all the intervention modules at school level. The school working group will therefore be responsible for the implementation of the following modules:

- Module 5: Education of children
- Module 6: Environmental changes related to physical activity (active playground)
- Module 7: Health-related physical education curricula
- Module 8: Environmental changes and school policy related to water consumption
- Module 9: Environmental changes and school policy related to fruit and vegetables

In the relevant modules, the task and responsibilities are explained in more detail.

3.1.3 Operation of the school working group

After setting up the school working groups, it is essential that local intervention teams start to talk with the school working groups as soon as possible in order to make a final decision on how the different modules will be implemented in the school.

As soon as the school working groups are set up, they have to act as a team. It must be avoided that the school project leader has to work out the whole project. For this reason a guideline has to be developed which supports the school project leader to implement the intervention. This guideline will include a well-defined description of the tasks of the school project leader, members and co-workers of the school working group, time table and description of the cooperation with local intervention teams andother useful information for the implementation of the intervention.

Operation of the school working group during the intervention adoption phase (Year 2)

During the first part of the adoption phase (September – June Year 2), the implementation of the intervention modules will be strongly supported by the local intervention team:

- All the participants of the school working group will receive instructions and guidelines about the modules that must be implemented in the schools/kindergartens.
- The school working group will work together with the local coordinator in order to implement the modules that will be imposed in the schools and kindergartens.
- The local coordinator has a coordinating role in the school working group and serves as the link to the community platform.
- The local coordinator and the school working group should meet on a regular basis and at least 4 times a year. Between these 4 meetings, the school working group can organise internal meetings whenever needed without the local coordinator.
- All these meetings make it possible to evaluate the implementation of the modules and to discuss and solve practical problems that have occurred or are expected.
- During the first meeting with the school working group, it is recommended to make an inventory of existing initiatives related to the prevention of obesity in the school or kindergarten. The following guidelines can be used during all the other meetings:
  - Produce a meeting report with the main points discussed and decisions taken. These reports can be used to analyse the implementation process of the intervention.
  - Discuss and evaluate the aspects of the intervention modules that have been executed as well as the ones forthcoming. Highlight things that went well (to be repeated), but also challenges (to find solutions) and things that went wrong (to try to avoid them happening again in the future). Provide strong support to the implementation team.
  - Continuously monitor activities going on inside the school or kindergarten related to preventing childhood obesity.
  - Do not ignore but take notice of new initiatives that are proposed by the school working group members.
  - Use the meeting as an opportunity to ensure the communication to all staff members of the school/kindergarten by publishing a newsletter or newsflash in the school paper.

The second part of the adoption phase (July – August, Year 2) is the transition period between the strongly supported and the supervised operation of the school working group. During the transition period, the following actions are recommended:

- The school working group should appoint a person in charge of the school working group (school working group coordinator) who will relay the local coordinator in his/her responsibilities with regard to the school working group.
- The school working group should start its activities at the beginning of September of year 3. The school working group coordinator should be invited to the community platform meetings.

Operation of the school working group during the intervention implementation phase (Year 3)

The school working group is expected to continue the activities and to work out new initiatives with minimal supervision and without continuous support of the local intervention team.

From September of Year 3 onwards, the local coordinator no longer participates in the meetings of the school working group. From that moment, the school working group coordinator should fully take over the tasks of the local coordinator.

The local coordinator must be kept informed about the initiatives of the school working group. For this reason, the local coordinator and the school working group or the school working group coordinator should meet at least twice during the intervention implementation phase, i.e. once between September and January of Year 3 and once between February and June of Year 3.

Between these meetings, if needed, the school working group can keep in touch with the local coordinator in order to solve practical problems or get advice.

Operation of the school working group during the intervention dissemination phase (September of Year 4 onwards)

The school working group is expected to operate completely independently, without any support or supervision from the local intervention team.

3.1.4 The creation of a school platform at community level

A school platform is a committee in which the school working groups of all the schools in the community can be represented.

Objectives

- Create a structure in which the schools can exchange knowledge and experiences, and start new initiatives for the prevention of childhood obesity.
- Enable all school working groups in the community to express a collective opinion and be considered as an important and full member of the community platform.

General recommendations

- For guidelines about the establishment and operation of the school platform, it is recommended to use the guidelines for the community platform (see module 1).
- The school platform should be the one being represented in the community platform, instead of all school working groups separately.

3.2 Module 5: Education of children

Objectives

- Integrate the key messages in the class curriculum.
- Increase knowledge, skills and self-efficacy in children.
- Promote key messages in the schools and kindergartens.

Concept of module 5

Module 5 includes the integration of the key messages in the class curriculum and the promotion of the key messages in the entire school. The key messages have to be integrated in the framework of (classroom) health education and not as part of obesity prevention to avoid stigmatisation of affected children.

To integrate the key messages in the class curriculum and to promote the key messages in the school, every participating school has to organise the **Healthy Weeks**: these are weeks in which a specific key message will be highlighted (with different exposures for primary schools and kindergartens).

Organisation of the Healthy Weeks during the adoption phase (Year 2)

- 9 Healthy Weeks should be organised during the intervention adoption phase, i.e. 4 Healthy Weeks about physical activity, 4 Healthy Weeks about diet and one additional Healthy Week about adequate sleep duration.
- The key message about spending time together (stress, coping and relaxation) will not be handled in a Healthy Week. This message should be systematically repeated and integrated within the other Healthy Weeks.
- As an example, Table 9 below shows how the Healthy Weeks can be planned during one school year. The planning of the Healthy Weeks can be culturally adapted, depending on the start of the school year and the local situation. However, it is recommended to maintain the alternation of the Healthy Weeks about physical activity and diet.

Intervention adoption phase (Year 2)		
Month of Year 2	**Theme of the Healthy Week**	**Key message to be highlighted**
Oct	Physical activity	Stimulating daily **physical activity**
Nov	Diet	Stimulating daily consumption of **fruit and vegetables**
Dec	Physical activity	Reduce **TV-viewing**
Jan	Diet	Stimulating the daily **consumption of water**
Feb	Physical activity	Stimulating daily **physical activity**
Mar	Diet	Stimulating daily consumption of **fruit and vegetables**
Apr	Physical activity	Reduce **TV-viewing**
May	Diet	Stimulating the daily **consumption of water**
Jun	Stress, coping and relaxation	Ensure adequate **sleep duration**
Jul	**Vacation period: No Healthy Weeks**	
Aug		

Table 9. Themes of the Healthy Weeks per month during the intervention adoption phase (Year 2)

The themes of the Healthy Weeks have to be synchronised with the themes of the window posters (module 2) and the themes of the educational folders (see later module 10). Table 10 shows how the Healthy Weeks can be synchronised with the window posters and the educational folders, respectively.

Intervention adoption phase (Year 2)			
	Window and street posters	Theme of the Healthy Weeks	Folders
Oct	Window poster related to physical activity and diet (1st poster)	Physical Activity	Distribution of folders related to physical activity and diet
Nov		Diet	
Dec		Physical Activity	
Jan		Diet	
Feb	Window poster related to physical activity and diet (2nd poster)	Physical Activity	Distribution of folders related to physical activity and diet
Mar		Diet	
Apr		Physical Activity	
May		Diet	
Jun	Window posters related to sleep duration (3rd poster)	Sleep duration	Distribution of folders related to sleep duration
Jul		Vacation period	
Aug			

Table 10. Synchronisation between the Healthy Weeks (module 5), the window posters (module 2) and the educational folders (module 10) during the intervention adoption phase (Year 2)

Tasks and responsibilities of the school working group and teachers to organise the Healthy Weeks

The school working group should:

- Display the window posters with the related key message in the school building.
- Communicate to the parents about the Healthy Weeks in the newsletter of the school.

The teachers should:

- Hang the window poster with the related key message in the classroom.
- Give the intervention package to the children: a folder (see module 10) and a window poster for the parents about the handled key message.
- Integrate the handled key message in the class lessons (instructions later on in this chapter).
- The exposure time of the Healthy Week in the classroom differs for the kindergarten and the primary school:

- Kindergarten: every day of the Healthy Week has to be in the theme of the handled key message.
- Primary school: teachers have to spend 1 class hour per Healthy Week for a total of 9 Healthy Weeks.

Tasks and responsibilities of the local intervention team

- Every school that participates in the intervention should be offered a manual for the organisation of the Healthy Week. This education manual should enable the schools and teachers to implement this module.
- The education manual has to be written by the local intervention team by following the guidelines described below.

Tasks and responsibilities of the school working group

- Distribution of the education manual to the relevant teachers.
- Communication to the teachers about the content of the education manual.
- Members of the school working group have to be approachable about practical problems related to this module (e.g. try to help the teachers with their problems or discuss them during the meetings).

Development of an education manual for the organisation of the Healthy weeks

Table 11 below provides possibilities and ideas that can be used for the development of the education manuals. These examples can count as a guideline but have to be worked out by the local intervention team, which implies an adaptation to the culture and the specific kindergarten and school structure in each country. It is not always necessary to develop new materials, and local intervention teams should also search for and select existing materials that can be included in the education manuals and that are able to fulfil the objectives of this module.

Practical guidelines to develop the education manuals

- The education manual has to contain educational strategies for the teachers to use in designing their lessons.
  - The ideas have to be concrete, original and age-adapted.
  - The information provided in the manual has to be ready to use.
- The education manual must provide the teacher with a practical answer to the question of how to handle the key messages in the classroom.
- When developing the education manual, always keep in mind that the educational strategies must aim to fulfil the following objectives for every key message:
  - Increase awareness and knowledge
  - Increase self-efficacy
  - Increase skills in children
- Because the content of the lessons will differ according to the age of the children, it is suggested to create a different education manual for the kindergarten and the primary school.
- Note: Education manuals are available for those countries where the IDEFICS intervention was tested for its effectiveness (Belgium, Cyprus, Estonia, Germany, Hungary, Italy, Spain and Sweden).

DAILY PHYSICAL ACTIVITY	
Awareness	- Interactive communication between teacher and children in combination with education material (posters, photographs, …): ▪ What does being physically active mean? ▪ What are the advantages of being physically active? ▪ When can you be physically active at home and at school? ▪ How much physical activity should we be getting each day? - Self report activities in class or for homework: children can fill in the physical activities they performed the day before on an attractive education worksheet, they can make a collage in which they present their physical activities or they can report their activities by assigning pictures (depending on the age). - Compare the self reports with the daily recommendation related to physical activity.
Skills	- Teach children to set physical activity goals. ▪ Communicate interactively (what are physical activity goals, what are the most important characteristics, …) and discuss some examples (case study). ▪ Provide educational worksheets on which children can report their physical activity goals at school and at home ("My accomplishment plan"). ▪ In a next step, the children have to report which activities they actually have done (review of the accomplishment plan). ▪ Performance against physical activity goals can be associated with a game: a child can gain a stamp or a sticker if they performed an activity goal. - Teach children how to motivate themselves to get physically active and how to maintain the motivation. For example: organise an interactive communication about "talking to yourself". ▪ Are the following sentences examples of positive or negative things that you can say to yourself? ▪ What are your feelings about …? ▪ How can you change negative sentences into positive sentences? - Play easy activity games in class (especially for toddlers and the youngest pre-school children) in order to teach locomotive skills.
Self-efficacy	- Help students to set realistic and challenging goals related to physical activity (see above). - Compliment children on the skills they have developed (positive feedback). - Encourage students to use effort as an explanation for failure, and the skills they have developed as an explanation for success. Do not attribute poor performance to lack of ability (and caution parents to avoid it as well). - Help children recognise the skills they are acquiring: make the children aware of what they have learned. - Success in the past is the best way to build confidence for future success: help children recognise their progress.

	▪ Let the children make an individual line graph of their progress. ▪ Give students frequent opportunities to update their progress. ▪ Provide a monthly calendar to record their daily progress. However, avoid class charts where individual progress is displayed because children will compare their progress in relationship to others. This will have a detrimental effect on the self-efficacy of those who do not progress as fast as others. Children have to learn to compare with themselves. - Use peer modelling during the lessons by working in small groups (this has an important influence on children's self-efficacy during skill acquisition).
REDUCE TV VIEWING	
Awareness	- Interactive communication between teacher and children: ▪ Which TV shows did you watch yesterday? ▪ When did you start watching? When did you stop? ▪ How many hours can we watch a day? Is it good to watch TV? - Self report activities in class or for homework: ▪ Encourage group reporting about time spent watching TV. ▪ Self report activities by means of an attractive education worksheet (did you watch TV while eating supper? Did you turn on the TV before school? Did you turn on the TV when you came home from school?, …). - Compare the self reports with the daily recommended limit of TV viewing.
Skills	- Have students brainstorm a list of fun alternatives to watching TV - Teach children how to watch TV selectively (for example: select the programmes you want to see and turn off the TV afterwards, help your parents in the kitchen during commercial breaks, …) - Teach self-monitoring techniques to children (for example: children make a diary in which they can report the amount of hours they watched TV, the programmes watched, …)
Self-efficacy	- See guidelines provided for daily physical activity and use them for the activities related to the key message about TV viewing.
DAILY CONSUMPTION OF FRUIT AND VEGETABLES	
Awareness	- Interactive communication between teacher and children in combination with education material (posters, photographs, …): ▪ What is good about eating fruits and vegetables? ▪ How many fruit and vegetables should we be eating each day? ▪ Where can we buy and eat fruits and vegetables? ▪ How can we eat fruits and vegetables? - Self report activities in class or for homework: children can fill in the amount of fruit and vegetables they ate the day before on an attractive education worksheet (or report it by making a collage or assigning pictures). - Compare the self reports with the daily intake recommendation for fruits and vegetables. - Visit a local fruit and vegetable farmer, grocery stores.

Skills	- Prepare fruits and vegetables in the class: show how to cut vegetables and how to peel fruit and provide the opportunity to practice in small groups (you can also provide an action plan that visualises the different steps in cutting a certain vegetable). - Organise a recipe competition in which children can use their preparation skills. - Organise recognition games.
Self-efficacy	- See guidelines provided for daily physical activity and use them for the activities related to the key message about fruit and vegetables. - As an extra, help children to serve as their own model. ▪ If a video camera is available, tape the children while they are participating in the preparation of fruit and vegetables and allow to view themselves being successful (or take photographs of the activity). While showing the tape or photographs, give positive feedback about the skills that the children acquired and were demonstrating.
Taste/preference	- Organise tasting activities and games in the class. - Provide fruits and vegetables that children have not tasted before on repeated occasions. - Teachers have to be role models: taste the fruits and vegetables in front of the children and show that you like them.
DAILY CONSUMPTION OF WATER	
Awareness	- Interactive communication between teacher and children in combination with education materials (posters, photographs, …): ▪ What are the advantages of drinking water? Why should I give preference to water over other drinks? ▪ How much water should we be drinking each day? ▪ Where can we drink water? In which forms can we drink water? - Self report activities in class or for homework: children can fill in the amount of water they drank the day before on an attractive education worksheet (or report it by making a collage or assigning pictures). - Compare the self reports with the recommended daily consumption of water.
Skills	- Show the children how to prepare flavoured water (e.g. with mint leaves, pieces of apple or strawberry). - Teach children how to deal with the daily recommended intakes (always fill your favourite cup with water and try to empty it while doing homework, always take a bottle of water with you, …).
Self-efficacy	- See guidelines provided for daily physical activity and use them for the activities related to the key message about the consumption of water.

Table 11. Ideas for the content of the education manuals

3.3 Module 6: Environmental changes related to physical activity - The Active Playground

Objectives

- Increase time spent in moderate to vigorous physical activities during recess.
- Provide an opportunity to help children reach the recommended physical activity level (contribute to the daily recommended norm of physical activity).

Possibilities to create an active playground

The creation and the promotion of an active playground is part of improving the school environment with the aim to stimulate the children to be physically active while at school.

The local intervention team has the opportunity to create an active playground in the schools and kindergartens by means of one or a combination of the following strategies:

Change the physical design of the playground

Redesign the playground, using multicolour playground markings (e.g. hopscotch) and physical structures (e.g. soccer goal posts, basketball hoops). This is found to be a sufficient stimulus for increasing children's school physical activity levels during recess and is also a method that is low in cost (Stratton & Mullan, 2005; Ridgers et al., 2007).

Provide attractive play tools in the playground

Providing game equipment during recess (e.g. balls, ropes, small bikes) is found to be effective in increasing children's physical activity levels. These findings suggest that promoting physical activity through game equipment provision during recess can contribute to reaching the daily recommended activity levels in children (Verstraete et al., 2006).

Structural changes related to recess period

The playground space available for children (number children/m^2) during recess periods is found to be an important predictor of children’s physical activity levels (Cardon et al.,2008). Therefore, structural changes that reduce the number of children on the same surface area can be an inexpensive way to increase physical activity levels during recess: e.g. divide all children in two groups and let them have playtime at different moments.

In addition to the environmental and structural possibilities to create an active playground, as described above, it is strongly recommended to combine one or more of these strategies with the following actions:

- Promotion of an active playground in the school.
- Coaching of teachers/educators to supervise the playground in an active way. They have an important role in stimulating children to be active in the playground.

Development of a manual for the creation of the active playground

Every school participating in the intervention should be offered a manual for the creation of an active playground. This school manual should enable the schools to implement this module. Because the content of the manual will differ according to the age of the children, it is suggested that a different manual be prepared for the kindergarten and the primary school. The school manual can be written by the local intervention team on the basis of the following guidelines:

Provide ideas to change the physical design of the playground

- Provide country-specific ideas to change the physical design of the playground.
- Integrate guidelines for the schools (e.g. if coloured marks are painted on the playground, they have to stay visible and thus have to be repainted on a regular basis).
- It is of special importance that the ideas are age-adapted: try to propose different ideas for the kindergarten and the primary school and even within the primary school. Castles, dragons, clock faces, mazes, fun trails, dens, hopscotch, letter scares, snakes and ladders, and various animals are popular in early primary schools. Markings for netball, football and short tennis, and targets for games-related skills are rather preferred in late primary schools (Stratton & Mullan, 2005). The ideas also have to reach both boys and girls.

Provide ideas for materials that can be provided in the playground

- Integrate different country-specific kinds of materials and play tools for the playground (e.g. sports balls or a (suit)case with circus materials).
- It is of special importance that the ideas are age-adapted (different for the kindergarten and the primary school) and that they reach both boys and girls.
- Integrate guidelines for the schools.

Provide suggestions on how the active playground can be promoted in the school

- Teachers have to inform the children about the possibilities to be active during break times (at the beginning of the school year and during the school year – e.g. by means of the school newspaper).
- The physical education teacher can instruct all students on proper use and all possibilities of the playground equipment.
- Just before the break, the teacher can remind the kids about the possibilities to be physically active.
- Ask the children after the break what they have done in the playground.

Include information about the importance of active supervision and coaching by teachers and/or educators

- Integrate guidelines for teachers/educators on how they should supervise/coach the children in the playground (e.g. encourage the students to walk when they talk instead of sitting on the bench, help the students to start a game or play along with them, and continue to encourage the students while they are being active and for other students to join in, challenge the children by setting goals e.g. how many times can you skip the rope in one minute, how many baskets can you score in one minute?).

Tasks and responsibilities of the school working group

- Distribution of the education manual to the relevant teachers.
- Communication to the teachers about the content of the education manual.

- Members of the school working group have to be approachable about practical problems related to this module (e.g. try to help the teachers with their problem or discuss them during the meetings).

3.4 Module 7: Health-related physical education curricula

Objectives

- Keeping all children active as much as possible during physical education classes (trying to reach a high activity level during the lessons).
- Development of knowledge, social attitudes and skills, and movement skills in children, which are necessary to lead an active lifestyle.
- Building children's confidence in their physical abilities.

Manual for the optimisation of physical education classes

- Every school participating in the intervention should be offered a manual which provides guidelines and tools for teachers, educators and nurses on how they can organise health-related physical education classes in primary school and how they can increase physical activity during time spent at the kindergarten.
- The manual should enable teachers, educators and nurses to implement this module.
- Because the content of the manual will differ according to the age of the children, it is suggested to create a different manual for the kindergarten and the primary school.
- The school manual can be written by the local intervention team on the basis of following guidelines (Bagby & Adams, 2007; SPARK, 2011):

Inform teachers, educators and/or nurses about the basic characteristics of health-related physical education curricula and emphasise them in the manual

- Aim to reach a high activity level for all the children during physical education classes.
- Develop the knowledge, attitudes, and social and movement skills, in children, which are necessary to lead an active lifestyle.
- Aspire to give every child positive experiences.
- Use activities with a high transfer value (i.e. activities children can also do in the playground and/or at home, e.g. rope skipping, Frisbee, …).

Integrate guidelines for teachers, educators and/or nurses to fulfil the basic characteristics of health-related physical education curricula

- Provide guidelines on how to increase the activity rate of children during physical education classes (e.g. restrict waiting time for children).
- Provide guidelines on which activities contribute to a higher physical activity level in children.
- Provide guidelines on how to create a pleasant lesson that stimulates the physical activity rate.

Provide strategies to integrate the guidelines in practice

- It is important that the teachers, educators or nurses verify which guidelines they are implementing already and how frequently. For example: a teacher, educator or nurse can focus on guidelines to reach a high activity level during one week. They verify which guidelines they already apply and which ones are new. The teacher, educator or nurse tries the new ones and evaluates the effect. Over the next weeks, the teacher, educator or nurse can then focus on other guidelines.

Note: Such a manual is available for those countries where the IDEFICS intervention has been tested for its effectiveness (Belgium, Cyprus, Estonia, Germany, Hungary, Italy, Spain and Sweden).

Tasks and responsibilities of the school working group

- Distribution of the education manual to the relevant teachers.
- Communication to the teachers about the content of the education manual.
- Members of the working group have to be approachable about practical problems related to this module (e.g. try to help the teachers with their problem or discuss them during the meetings).

3.5 Module 8: Environmental changes and school policy related to water consumption

Objectives

- Create a school environment that discourages the consumption of sugar-sweetened beverages.
- Create a school environment that stimulates the consumption of water.

Possibilities to increase the daily water consumption in schools and kindergartens

The local intervention team can increase the daily water consumption in schools and kindergartens by means of one or a combination of the following strategies:

Permanently provide free water during breaks, play time and/or lessons

- Provide water fountains in the playground (environment).
- Provide free water at the table during dinner (environment).
- Allow drinking water during theory lessons: children can have a water bottle on their desk (school policy).
- Allow drinking water during physical education classes and stimulate the teacher to integrate a "water drinking moment" (school policy).

It is of special importance that the provision of water facilities is in proportion to the number of children in the school and adapted to the student population (location, height, hygiene, …). It is for example better to allow drinking cups in the kindergarten whereas the placing of water fountains is rather recommended in primary schools. The provision of water should also be clearly communicated to the students and the staff members: when, where, how, …

Reorientation of the beverage supply in the school (changing the environment and/or the school policy)

Favour water as the main drink supply in school. This reorientation will require clear regulations about the consumption of sugar-sweetened beverages and water and possibly a change to the school policy. If it is possible to change the school policy or to set regulations about the consumption of sugar-sweetened beverages, this should be clearly communicated to the children, staff members and parents.

Examples of how the beverages supply can be changed:

- Remove vending machines in which sugar-sweetened beverages are offered (environment – school policy).
- Do not remove all vending machines but replace sugar-sweetened beverages by water or other healthier options (environment – school policy).
- Do not allow the sale and consumption of sugar-sweetened beverages in the school (regulations – school policy)

Promote the availability of water and clearly communicate about the regulations

It is strongly recommended to promote all the initiatives to the children, parents and staff members of the school. It is also essential to communicate to the children, parents and staff members about the regulations related to the consumption of water. Examples of how the availability of water can be promoted:

- Describe the water initiatives and regulations in the school paper.
- Teachers can ask children to make an advertisement about the consumption of water and make them visible in the school.
- Provide recyclable cups or tins for water (sponsorship can be used to finance this initiative).

3.6 Module 9: Environmental changes and school policy related to fruit and vegetables

Objective

- At least once a week, make fruit and vegetables available in the school environment during the breaks.

Possibilities to increase the consumption of fruit and vegetable in schools and kindergartens

The local intervention team can increase the daily consumption of fruit and vegetables in schools and kindergartens by means of one or a combination of the following strategies:

Availability and accessibility of fruit and vegetables in the school (make environmental changes)

The working group of the school can make a contract with a local fruit and vegetable trader or merchant who can deliver at least once a week fresh and seasonable fruit and/or vegetables to the school.

Practical implications

- Try to provide fresh seasonal fruit.
- Try to provide fruit that is easy to eat for the children.
- Make the fruits and vegetables accessible, especially for the youngest children: involve for example volunteering (grand)parents to prepare the fruits and vegetables at school.
- Sponsorship can be used to finance the fruit and vegetable project.
- Without sponsorship, a child can participate in the fruit and vegetable project when parents have paid a financial contribution to the school at the beginning of the school year.

- The day on which the fruits and vegetables are delivered can become a dedicated 'fruit and vegetable day' for the rest of the school year.

Other possibilities

- Organise: "bring your own fruits and vegetables to school".
- Organise an additional "fruit and vegetable day": children can only bring fruit and/or vegetables instead of less healthy snacks.
- School gardens (children grow vegetables at school).
- Organise the calendar game: every child that brings a piece of fruit or gets a sticker in their agenda.

Reorientation of the food supply in the schools (changing the environment and/or the school policy)

Change the food supply in the school to include fruits and vegetables as an alternative to less healthy snacks. This reorientation will require clear regulations and possibly a change to the school policy. If it is possible to change the school policy or to set regulations about the consumption of less healthy snacks and fruits and vegetables, this should be clearly communicated to the children, staff members and parents. Examples of how the food supply can be changed:

- Remove the vending machines in which less healthy snacks are represented.
- Restrict the time that vending machines with (less healthy) snacks are available.
- Increase the price of high-energy snacks and lower the price of fruits and vegetables.
- Do not sell less healthy snacks but replace them by fruit and/or vegetables.

Promote the availability/accessibility of fruit and vegetables and clearly communicate about the regulations

It is strongly recommended to promote all the initiatives to the children, parents and staff members of the school. It is also essential to communicate to the children, parents and staff members about the regulations related to the consumption of water. Examples of how the availability/accessibility of fruit and vegetables can be promoted:

- Describe the initiatives and regulations related to fruits and vegetables in the school newspaper.
- Inform the parents in particular about the fruit and vegetables initiatives and try to involve them.
- Teachers can ask children to make an advertisement about the consumption of fruit and vegetables and make them visible in the school.

4. IDEFICS intervention modules at family level

4.1 Module 10: Education of parents

Objectives

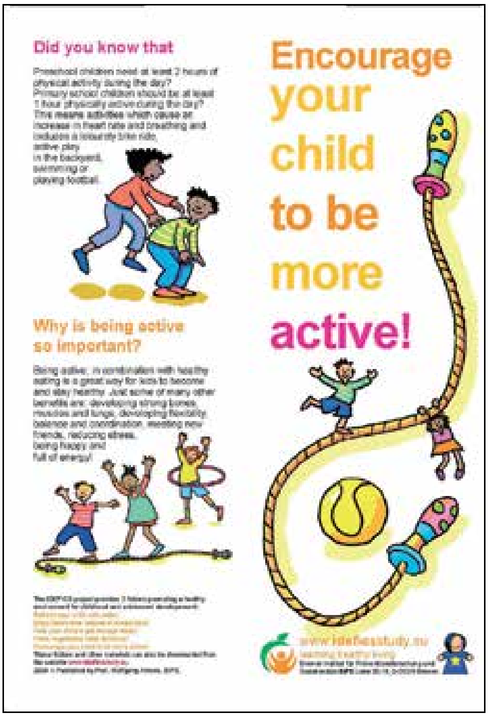
- Increase behavioural skills in parents in order to increase social support and accessibility and availability of fruit and vegetables at home.
- Increase awareness in parents.
- Increase self-efficacy in parents.

Concept of module 10: educational folders and videos for parents

Parents should receive educational folders and videos to learn about parenting strategies that can remove barriers and facilitate them in their ability to create health-promoting family environments.

Educational folders

The local intervention team can develop 3 folders: 1 folder about diet, 1 folder about physical activity and 1 folder about sleep duration. The key message about spending more time together has to be integrated in these folders. The educational folder about physical activity that was used in the IDEFICS intervention, is displayed below. All of the folders are already available in the languages of those countries where the IDEFICS intervention was tested for its effectiveness (Belgium, Cyprus, Estonia, Germany, Hungary, Italy, Spain and Sweden).



Did you know that

Why is being active so important?

Encourage your child to be more active!

www.idefics study.eu

5 tips to have your kids to be more active and to live a more healthy lifestyle

1. You can spend time together with your child and be an active parent

2. You and your child can try to be active any way you can

3. You can stimulate and support your child to be active

4. You can reduce inactive leisure time and replace it by active alternatives

5. You can support and encourage your child's active choices

Guidelines for the development of the folders

Four aspects of parenting can be the focus of the educational folders and can provide a framework for the development of the educational materials, i.e. beliefs and knowledge of the parents, parental modelling, availability and/or accessibility and shaping. For a detailed description of these aspects of parenting, references are made to Crawford and Jeffery (2005). Based on the examples provided by Crawford and Jeffery (2005), Table 12 offers an overview of strategies that can help parents in their ability to create family environments that promote and encourage a healthy lifestyle.

DAILY PHYSICAL ACTIVITY	
Beliefs and knowledge	- Use the daily recommended norm related to physical activity as a guideline. - Believe in the ability of your children to be physically active.
Modelling	- Be physically active yourself and together with the family. - Include physical activity into the leisure time of the family (e.g. hiking or cycling with the entire family). - Plan and participate in physically active family activities (e.g. walking or cycling instead of driving, playing outside) and include physical activity in family events such as birthday parties, picnics, and vacations.
Availability and/or accessibility	- Provide activity-related equipment in the home environment (e.g. balls, bicycles). - Visit sports and recreation facilities with your children where they can be physically active (e.g. community playgrounds, sports organisations). - Identify outdoor activities and perform them together with your child. - Identify indoor activities other than screen-based activities and perform them together with your child.
Shaping	- Use fun physical activities as a reward for behaviours with positive outcome. - Do not use physical activity as a punishment for behaviours with negative outcome.
REDUCE TV VIEWING	
Beliefs and knowledge	- Use the daily recommended norm related to watching television as a guideline.
Modelling	- Reduce your own TV viewing time. - Switch off the TV while you, your children or the entire family is eating.
Availability and/or accessibility	- Set clear rules regarding TV viewing time (e.g. your children can select one or more programmes, they can watch these programmes but the TV should be switched off before and after the selected programmes). - Do not put a TV or a computer in your children's bedroom. - Provide active and fun alternatives to TV viewing.
Shaping	- Do not use watching TV as a reward for behaviours with positive outcome.

DAILY CONSUMPTION OF WATER	
Beliefs and knowledge	- Use the daily recommended norm related to water as a guideline.
Modelling	- Limit your own intake of less healthy beverages and model the consumption of water at every moment of the day. - Show your child that you like water and that water is tasty. - Do not show your own dislike for water.
Availability and/or accessibility	- Provide water during meals. - Provide sometimes alternatives for clear water, for example flavoured water with mint leaves or pieces of strawberry, apple, … - Provide your child with a nice tin or their favourite cup which is always filled with water and accessible to take. - Set rules about the consumption of less healthy beverages: e.g. only one can of soft drink a week, only on weekends, only at parties, …
Shaping	- Do not use sugar-sweetened beverages or similar drinks as a reward for behaviours with positive outcome. - Do not use drinking water instead of sugar-sweetened beverages as a punishment for a behaviour with negative outcomes.
DAILY CONSUMPTION OF FRUIT AND VEGETABLES	
Beliefs and knowledge	- Use the daily recommended norm related to fruit and vegetables as a guideline.
Modelling	- Eat fruits and vegetables yourself and show your children that you like them. - Involve your children in the selection and preparation of fruit and vegetables.
Availability and/or accessibility	- Make fruits and vegetables easy available and accessible in the home environment and during family trips. - Provide fruits and vegetables that your child never has tasted before. - Provide fruit and vegetables in a form that is easy and ready to eat (e.g. pre-cut vegetables and/or peeled fruit). Try to do this in the home environment, during family trips and/or to provide it to your child as a snack at school (instead of less healthy snacks).
Shaping	- Do not use fruits and vegetables as a punishment for behaviours with negative outcome. - Do not use unhealthy snacks and foods as a reward for behaviours with positive outcome.

Table 12. Examples of parenting strategies for diet and physical activity (Crawford & Jeffery, 2005)

Distribution of the folders

- The community platform is responsible for the dissemination of the folders at the right moment in all the different sectors of the community (see stakeholders). For example: the school is an important and easy setting to distribute the folders to the residents of the community. Thus, the community platform also has to take care that the schools receive the folders in time.
- The distribution of these folders will be synchronised with the distribution of the window and street posters and the integration of the key messages in the class curriculum (see Table 10): use the folders that handle the same key message as highlighted on the window poster and in the Healthy Weeks.

Educational videos

A media agency could develop scenarios for the educational videos based on the content and ideas provided in this intervention manual.

5. Conclusion

The fundamental idea of the IDEFICS project was that obesity prevention should be triggered by local policy makers or local public health authorities and supported by initiatives taken by local stakeholders in the community. As part of this project, a community-based intervention for the prevention of childhood obesity was developed and tested for its effectiveness. The content and the implementation strategy of the IDEFICS intervention are now available and provide local policy makers and public health authorities with the opportunity to explore a model for obesity prevention in Europe.

This chapter was written in a way that local policy makers or local public health authorities have the necessary guidance and recommendations at their disposal to implement the intervention in their local city or community. The major advantage of the intervention framework is that it can be culturally adapted to the local needs and requirements which increases the feasibility of implementation in European countries.

6. Acknowledgements

This work was done as part of the IDEFICS Study (www.idefics.eu). We gratefully acknowledge the financial support of the European Community within the Sixth RTD Framework Programme Contract No. 016181 (FOOD).

7. References

Ahrens, W., Bammann, K., Siani, A., Buchecker, K., De Henauw, S., Iacoviello, L., Hebestreit, A., Krogh, Vittorio, L., Lauren, M., Staffan, Molnar, D., Moreno, L.A., Pitsiladis, Y., Reisch, L.A., Tornaritis, M., Veidebaum, T., Pigeot, I., on behalf of the IDEFICS consortium (2011). The IDEFICS cohort: Design, characteristics and participation in the baseline survey. *Int J Obes, 35* (Suppl.1):S3 - S15.

Ahrens, W., Bammann, K., De Henauw, S., Halford, J., Palou, A., Pigeot, I., Siani, A., Sjöström, M. (2006). Understanding and preventing childhood obesity and related disorders–IDEFICS: a European multilevel epidemiological approach. *Nutr Metab Cardiovasc Dis, 16*, 302-308.

Bagby, K. & Adams, S. (2007). Evidence-based practice guideline: increasing physical activity in schools--kindergarten through 8th grade. *J Sch Nurs, 23,* 137-143.

Cardon, G., Van Cauwenberghe, E., Labarque, V., Haerens, L., & De Bourdeaudhuij, I. (2008). The contribution of preschool playground factors in explaining children's physical activity during recess. *Int J Behav Nutr Phys Act, 26*, 5:11.

Crawford, D. & Jeffery, R. W. (2005). Obesity Prevention and Public Health. Oxford University Press.

Davison, K.K. & Birch, L.L. (2001). Childhood overweight: a contextual model and recommendations for future research. *Obes Rev*, 2, 3.

De Henauw, S., Verbestel, V., Marild, S., Barba, G., Bammann, K., Eiben, G., Hebestreit, A., Iacoviello, L., Gallois, K., Konstabel, K., Kovacs, É., Maes, L., Molnar, D., Moreno, L.A., Reisch, L., Siani, A., Tornaritis, M., Williams, G., Ahrens, W., De Bourdeaudhuij, I., Pigeot, I. (2011). The IDEFICS Community-Oriented Intervention Program: A new model for childhood obesity prevention in Europe? *Int J Obes (Lond), 35 Suppl 1*:S 16-23.

Ferreira, I., van der, H. K., Wendel-Vos, W., Kremers, S., van Lenthe, F. J., & Brug, J. (2007). Environmental correlates of physical activity in youth - a review and update. *Obes Rev, 8*, 129-154.

Pigeot, I., De Henauw, S., Foraita, R., Jahn, I., Ahrens, W. (2010). Primary prevention from the epidemiology perspective: Three examples from the practice. *BMC Med Res Meth, 10* (10).

Ridgers, N. D., Stratton, G., Fairclough, S. J., & Twisk, J. W. (2007). Long-term effects of a playground markings and physical structures on children's recess physical activity levels. *Prev Med., 44*, 393-397.

Sports, Play and Activity Recreation for Kids (SPARK) (2011). Pre-developed curricula specifically designed to increase levels of physical activity in children during physical education class.
http://www.sparkpe.org/

Stratton, G. & Mullan, E. (2005). The effect of multicolor playground markings on children's physical activity level during recess. *Prev Med., 41*, 828-833.

Verbestel, V., De Henauw, S, Maes, L, Haerens, L., Mårild, S., Eiben, G., Lissner, L., Moreno, L.A., Lascorz Frauca, N., Barba, G., Kovács, É., Konstabel, K., Tornaritis, M., Gallois, K., Hassel, H., De Bourdeaudhuij, I. (2011). Using the intervention mapping protocol to develop a community-based intervention for the prevention of childhood obesity in a multi-centre European project: the IDEFICS intervention. *Int J Behav Nutr Phys Act, 8*, 82.

Verstraete, S. J., Cardon, G. M., De Clercq, D. L., & De Bourdeaudhuij, I (2006). Increasing children's physical activity levels during recess periods in elementary schools: the effects of providing game equipment. *Eur J Public Health, 16*, 415-419.

2

Testing the Assumptions of Stage of Change for Fruit and Vegetable Consumption: A Naturalistic Study

Jay E. Maddock*, Jodi D. Barnett, Carrie S. Marshall and Claudio R. Nigg
Department of Public Health Sciences
University of Hawai'i at Mānoa,
USA

1. Introduction

Chronic disease now accounts for 7 of every 10 deaths in the United States and 60% of the nation's health expenditures [1]. Poor nutrition is a substantial contributor to the chronic disease burden, accounting for over $33 billion in medical costs and $9 billion in lost productivity per year [1]. Fortunately, many detriments of chronic disease, such as increased risk for heart disease [2-4], stroke [5, 6], diabetes [7, 8], osteoporosis [9], and cancer [10, 11], can be prevented through adoption of a healthy diet. Fruit and vegetables are an integral part of a healthy diet, and they provide many nutrients that may reduce the risk for some types of cancers and chronic disease [12-16]. To achieve this protective effect, disease prevention guidelines recommend that individuals consume at least five servings of fruits and vegetables a day [17, 18]. However, data from the 50 US states indicates that 70-80% of US adults fall short of these recommendations [19].

The substantiated link between poor diet and the epidemic prevalence of chronic disease in America necessitates population-based interventions aimed at increasing fruit and vegetable consumption. Explanatory theories of behavior change, such as the Transtheoretical Model (TTM), can help guide intervention programs in developing the most effective strategies for promoting and sustaining change in a population. Over the past two decades, the central organizing construct of the TTM, the stages of change has experienced widespread use as well as pointed criticisms [20]. The model postulates that people move through a series of five stages of change in their attempts to modify their problem behaviors [21]. As people change stages, they employ mediating processes such as self-efficacy and decisional balance, differentially making each stage unique. Five stages of change have been identified: precontemplation (no intention to change behavior in the foreseeable future, or denial of need to change); contemplation (intention to change within the next 6 months); preparation (serious intention to change in the next 30 days); action (initiation of overt behavioral change); and maintenance (sustaining behavioral change for 6 months or more) [21].

---

* Corresponding Author

Although the TTM was originally developed as a framework for smoking cessation, it has gained widespread use over the past two decades as the basis of formalized treatment programs and population-based interventions for over a dozen health-related behaviors. The TTM has been adapted to many areas of behavior change including eating behaviors [22], exercise adoption [23], condom use [24], and several others [25].

With the great popularity of the stages of change construct in particular, careful measurement work has not always been conducted when adapting the construct to a new behavior. Over the past few years, several research papers have examined the cross-sectional relationship between the stages of change for fruit and vegetable intake and other related variables [26-31].

Research on self-change in naturalistic populations is necessary to assess the efficacy of stage of change models. While the vast majority of the literature is cross-sectional, a longitudinal approach is "more in line with the temporal nature of the model" because it can expose processes and patterns of change at the individual level that may be masked by a cross-sectional, population-based design [32].

Weinstein and colleagues (1998) have outlined four properties of a stage theory of health behavior [33]. The first is a classification system to place individuals into discrete stages. The second characteristic is an ordering of the stages. It is assumed here that although people can move both forward and backward between the stages they are most likely to move to adjacent stages in their attempts to change. It is also predicted that on a population level the closer a stage is to action, the more likely those people are to move into action in the future. The final two characteristics of a stage theory are common barriers to change facing people in the same stage and different barriers to change in different stages. For instance, the TTM postulates that experiential processes are important for early stage changes such as precontemplation to contemplation, while behavioral processes are important for later stage changes for example preparation to action [21]. Two alternatives also exist to the interpretation of a stage model: pseudostages created from a continuous variable, for instance motivation (linear pseudostage), and pseudostages created from a general algebraic equation, including interactions and limits on variables (non-linear pseudostage) [33].

Four research designs have been developed for testing the efficacy of stage models [33]. The first and most common design is to examine cross-sectional comparisons of people in different stages. In this approach, an analysis of variance is typically conducted to assess differences across the stages of change for certain variables which are predicted to differ by stage. While the largest body of research is conducted in this area it provides a weak test of stage model which cannot rule out a non-linear pseudostage [33]. The second design for testing the stage of change construct is the examination of stage sequences. This requires longitudinal data and often predicts movement between pre-action stages to action [34]. The prediction is that people who start in stages closer to action will be more likely to move to action over time. Stage transitions across more than two time points can also be examined to assess if changes are more likely to occur to adjacent stages. However, data collection periods can often miss transitions in stage. This also does not rule out either pseudostage. The third design is the longitudinal prediction of stage transitions. This design tests the assumption that different constructs are important for different stage transitions. For instance, behavioral processes are more important for the transition from preparation to action than from the transition from precontemplation to contemplation. This data helps establish that motivation is not a continuum but that real quantitative differences occur between the stages, supporting the stage model and the non-linear pseudostage model [33]. The final design is an experimental study of matched and mismatched interventions, where participants are randomized to either a stage appropriate intervention or a non-stage appropriate intervention. This method provides the best test of the stage model by formally testing the assumptions that a stage matched intervention is superior to a mismatched intervention [33].

In this study, we will examine three of Weinstein and colleagues (1998) tests of a stage model [33], summarized in Table 1. First, cross-sectional comparisons of people in different stages will be assessed by behavior and related psychosocial variables. Then stage sequences will be examined over three time points. Finally, longitudinal prediction of stage transitions by baseline behaviors and related psychosocial variables will be conducted.

1. Cross-sectional comparisons of people in different stages Supports a stage theory if: a. Attributes of people differ across stages b. The patterns of differences across stages vary from one attribute to another
2. Examination of Stage Sequences Supports a stage theory if: a. Successive stages follow the hypothesized sequence
3. Longitudinal Predictions of Stage Transitions Supports a stage theory if: a. Predictors of stage transition vary from stage to stage

Table 1. Research Designs for Testing Stage Theories

2. Methods

2.1 Data collection

A longitudinal survey using random digit dialing of Hawaii's non-institutionalized adult population was conducted from February to April of 2002 [34]. The person over 18 who had the last birthday was asked to complete the interview to provide randomization within household. Informed consent was obtained over the phone. The survey took approximately 20 minutes to complete. All procedures were approved by the University of Hawaii Committee on Human Subjects. Interviewers were trained on the survey in small group settings for 6 hours in both classroom and live phone settings. Interviewers were assisted by a computer aided telephone interview (CATI) system designed specifically for the survey. Skip patterns and out of range responses were automatically controlled by the system. Follow-up surveys were conducted at 6 and 12 months post-baseline.

During the follow-up, at least five attempts were made to contact the participants. Respondents were also given the option to callback at their convenience. Primary reasons for non-completion included disconnected phone numbers, no longer living at current number, no callback by participants, and no answer.

2.2 Measures

Participants were asked a series of demographic questions, including age, sex, height, weight, education attained, income level, marital status, ethnic identification, language spoke at home, and perceived health. Participants were then asked about behaviors and other variables related to fruit and vegetable consumption.

Fruit and Vegetable Intake was assessed using a short "all day" assessment developed by the National Cancer Institute [35]. This instrument has been shown to have good reliability compared to actual intake and is recommended for population based research [35].

Stage of Change relative to consumption of fruits and vegetables was assessed. The instrument (Figure 1) inquired about participants' fruit and vegetable intake followed by their intentions to consume five or more servings per day. Participants were classified into one of five stages; (1) Precontemplation – Do not eat 5-a-day with no intentions to do so in the next 6 months; (2) Contemplation – Do not eat 5-a-day but intend to do so in the next six months; (3) Preparation – Do not eat 5-a-day but intend to in the next month; (4) Action – Currently eating 5-a-day, but for less than six months; (5) Maintenance - Currently eating 5-a-day for more than six months [36].

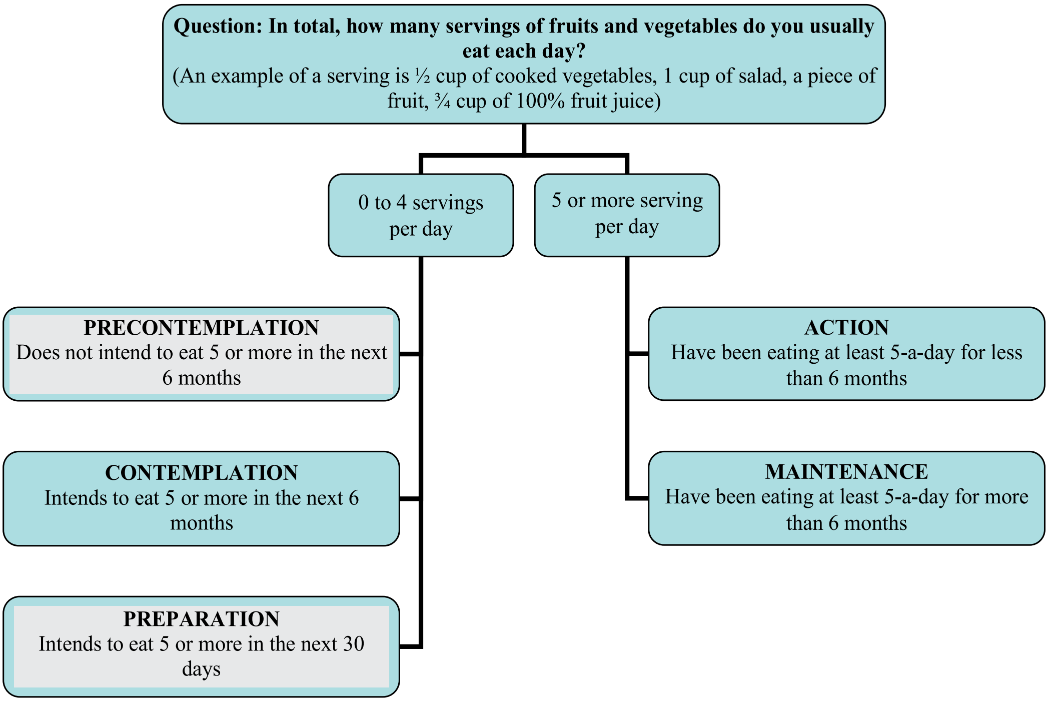


Fig. 1. Stage of Change Instrument

Self-Efficacy for eating 5-a-day was assessed with a 10-point continuous scale from (1) Not at all confident to (10) Very confident; "How confident are you that you can eat 5 or more servings of fruits and vegetables per day?"

Intentions, subjective norm, perceived behavioral control and attitude from the Theory of Planned Behavior were also assessed [37].

Subjective Norm was measured using three continuous 10-point scale items from (1) strongly disagree to (10) strongly agree. Satisfactory alpha levels for this scale were α = .82 and .86 at baseline and 1 year. Questions included: "Most people who are important to me think I should be eating 5-a-day on a regular basis"; "People in general approve of eating 5-a-day"; and "Most people who are important to me would like me to eat 5-a-day".

Perceived Behavioral Control was measured using three continuous 10-point scale items. The higher the score meant the greater the endorsement of the item. Satisfactory alpha levels for this scale were α = .76 and .79 at baseline and 1 year. Questions included: "If I wanted to I could easily eat 5-a-day on a regular basis" (1) strongly disagree to (10) strongly agree; "How much control do you have over the number of times you eat 5-a-day" (1) very little control to (10) complete control; and "For you to eat 5-a-day on a regular basis is……" (1) extremely difficult to (10) extremely easy.

Attitudes were measured using four continuous 10-point scale items. Satisfactory alpha levels for this scale were α = .83 and .84 at baseline and 1 year. One question was asked followed by four different response sets: "Eating 5-a-day regularly is…." (1) harmful to (10) beneficial; (1) foolish to (10) wise; (1) unenjoyable to (10) enjoyable; and (1) unpleasant to (10) pleasant.

2.3 Analysis

The SPSS System for Windows version 11.5.0 was used for all data analyses [38]. For all analyses, only individuals who completed both the baseline and 12 month assessment were included. T-tests and chi-square analyses were used to assess differential dropout between baseline and the 12 month follow-up. Mean differences in behavior, self-efficacy, intentions, subjective norm, perceived behavioral control and attitudes for fruit and vegetable consumption were examined across stages of change at baseline using an analysis of variance (ANOVA). Longitudinal stability of the stages of change was examined using the Cohen's kappa test of inter-rater reliability. Stage transitions were examined descriptively by examining the pattern of stage transition across baseline, 6 months and 12 months. For this analysis, the contemplation and preparation stages were combined as were the action and maintenance stages. This was done due to very small sample sizes in both contemplation and action. Patterns of change were assessed using similar categories to Prochaska and colleagues (1991). They included: 1. Stable – in the same stage all three time points; 2. Progressing – moving forward at least one stage with no setbacks; 3. Relapse – moving backward at least one stage without returning to the original stage; 4. Mixed inverted V pattern – participants first increased and then decreased their stage, for example moving from contemplation to action and then back to contemplation; 5. Mixed V pattern – participants first decreased and then increased their stage, for example moving from action to contemplation and then back to action [39].

Baseline differences in behavior and the related psychosocial variables were examined as possible stage change predictors by looking at each baseline stage of change and assessing which stage the individual moved into at 12 months. One-way ANOVAs with follow-up Tukey tests were used to assess significant differences by baseline stage.

3. Results

3.1 Sample characteristics

A total sample of n=3,519 respondents completed baseline surveys (22.8% response rate), with n=2,390 (67.9%) completing the six month follow-up, and N=1,978 completing the 12 month follow-up (56.2% of baseline). At baseline, the mean age was 46.3 years, with a mean of 14.9 years of education, and a median income of $40,000 to $50,000 per year. A greater percent of females participated in the study (62.0%). The sample was ethnically diverse with 86% of the sample comprised of Caucasian, Japanese, Filipino, Chinese, and Native Hawaiians. A mean of 3.1 fruits and vegetables were consumed per day by participants. All demographic variables are displayed in Table 2.

3.2 Attrition analysis

Differential attrition rates by demographic variables and behavior were analyzed. No significant differences were found across gender and fruit and vegetable consumption. Small but significant differences existed across age, income level, education, and ethnicity, $p < .001$. Those less than 35 years of age, a household income of less than $40,000, and no college education experienced a higher rate of attrition. Those of Japanese ethnicity showed the greatest percent of retention as compared to other ethnicities. The main reason for non-response was an inability to contact the participant (63.3%) after five attempts. Demographic characteristics of participants who completed the study and those who dropped out are presented in Table 2.

3.3 Missing data analysis

Missing data was assessed for all of the variables of interest including stage of change, fruit and vegetable consumption, intentions, self efficacy, attitudes, subjective norms and perceived behavioral control. None of these variables had more than 10% missing data across both baselines, 6 months and 12 months. With the low level of missing data and the large sample size, no corrections were made for missing data and the case was eliminated from the analysis.

3.4 Stage distributions

First, the stage distributions were examined across both time points, individually. Almost identical results were found with 39.2% in precontemplation (38.2% at 12 months), 5.4% in contemplation (4.2% at 12 months), 34.8% in preparation (36.7% at 12 months), 2.2% in action (1.9% at 12 months) and 18.4% in maintenance at baseline (18.9% at 12 months).

	6 Months		12 Months	
Demographics	Dropout (n = 1129)	Retain (n = 2390)	Dropout (n=1541)	Retain (n=1978)
Gender				
% Male	41.4	38.9	41.9	38.0
% Female	58.6	61.1	58.1	62.0
Age				
Mean (SD) in yrs	42.8 (16.4)	**48.0 (16.0)	43.1 (16.4)	**48.8 (15.8)
Ethnicity[1,2]				
% Caucasian	32.5	35.0	33.2	35.0
% Hawaiian / part-Hawaiian	18.7	17.6	19.3	16.9
% Japanese	15.5	20.7	14.9	22.2
% Filipino	10.8	8.7	11.0	8.2
% Other	22.5	18.0	21.7	17.7
Education				
Mean yrs (SD)	14.2 (3.0)	**14.9 (3.1)	14.4 (3.0)	**14.9 (3.2)
Income[1,2]				
< $40,000	49.8	39.3	47.4	39.0
$40,000-$49,999	11.9	11.2	12.2	10.9
> $49,999	38.3	49.5	40.4	50.1
Behavior				
Mean fruit and vegetable intake	3.0 (1.83)	3.1 (1.82)	3.1 (1.9)	3.1 (1.8)

** Significant t-test (p < .001)
1 = Significant Chi-Square Test at 6 months (p < .001)
2 = Significant Chi-Square Test at 12 months (p < .001)

Table 2. Attrition comparison by baseline demographics at 12 months

3.5 Behaviors and related constructs by stages of change

Differences in behavior and related constructs for fruit and vegetable consumption were examined across stages of change at baseline for self-efficacy, intention, subjective norms, perceived behavioral control, and attitude. Significant differences in behavior and all related constructs were observed across stages of change, $p < .001$ (Table 3). The largest effect sizes were found across stages of change for behavior, intentions and self efficacy; $\eta^2 = .624$,

η^2=.310 and η^2=.304 respectively. These three variables showed significant differences across almost all of the stage transitions. Attitudes and subjective norms were important for differentiating between precontemplation and the other stages and perceived behavioral control was significant in differentiating between maintenance and the other stages.

	Behavior	**Self-Efficacy**	**Attitudes**	**Subjective Norm**	**Perceived Behavioral Control**
	N = 1884	N = 1917	N = 1890	N = 1791	N = 1891
Precontemplation	2.09 (1.02)	5.32 (2.92)	8.23 (1.83)	6.45 (2.64)	6.69 (2.30)
Contemplation	2.51 (1.14)	6.65 (2.52)	8.57 (1.68)	7.40 (2.30)	7.22 (2.23)
Preparation	2.73 (1.00)	7.94 (2.00)	9.02 (1.47)	7.82 (2.34)	7.97 (1.86)
Action	5.53 (0.91)	8.19 (1.76)	9.23 (1.08)	8.02 (1.87)	7.91 (2.01)
Maintenance	5.87 (1.44)	9.27 (1.37)	9.55 (0.98)	8.34 (1.95)	9.20 (1.19)
F-value	F(4,1879) = 780.39, p < .000	F(4,1912) = 209.03, p < .000	F(4,1885) = 49.97, p < .000	F(4,1786) = 45.50, p < .000	F(4,1886) = 103.97, p < .000
eta^2	.624	.304	.096	.092	.181
Tukey post-hoc test[1]	PC < All C < A,M P < A,M	PC < All C < P,A,M P,A < M	PC< P,A,M C < M	PC < All C < M P < M	PC< P,A,M C < P,M P,A < M

PC = Precontemplation, C = Contemplation, P = Preparation, A = Action, M = Maintenance.

Table 3. Behavior and mediators by stage of change at baseline

3.6 Examination of progression to action/maintenance

While a high level of stability in stages of change was found cross-sectionally, longitudinal results yielded a small yet significant Cohen's kappa correlation at $\kappa = .246$, indicating a low level of stability from baseline to 12 months, $p < .001$ (Table 4). Less than half of those in maintenance at baseline remained after 12 months. Precontemplators showed the greatest level of stability at 59.2%. Baseline stage of change was a significant predictor of action and maintenance status at 1 year with only 8.0% of precontemplators reaching action or maintenance compared to 7.9% of contemplators, 22.1% of individuals in preparation, 38.1% of individuals in action and 49.1% of individuals in maintenance (p < .01). Table 4 displays the stage transitions from baseline to 12 months.

	Stage at 6 Months n = 2390				
Baseline[1]	PC (%)	C (%)	P (%)	A (%)	M (%)
Precontemplation	**58.8**	4.9	27.4	2.4	6.6
Contemplation	36.0	**16.2**	33.3	3.6	10.8
Preparation	25.3	4.8	**49.6**	2.1	18.2
Action	14.6	2.1	41.7	**2.1**	39.6
Maintenance	15.3	1.1	26.5	4.6	**52.5**
	Stage at 12 months n = 1978				
Baseline[2]	PC (%)	C (%)	P (%)	A (%)	M (%)
Precontemplation	**59.2**	5.1	27.7	0.8	7.2
Contemplation	37.3	**8.8**	46.1	2.0	5.9
Preparation	26.9	3.7	**47.3**	2.6	19.5
Action	14.3	7.1	40.5	**4.8**	33.3
Maintenance	18.4	1.7	30.7	2.3	**46.8**
	Stage at 12 months n = 1831				
6 Months[3]	PC (%)	C (%)	P (%)	A (%)	M (%)
Precontemplation	**66.3**	5.1	22.6	0.2	5.9
Contemplation	34.5	**8.3**	48.8	4.8	3.6
Preparation	26.5	2.7	**54.0**	2.7	14.1
Action	24.5	8.2	32.7	**2.0**	32.7
Maintenance	11.9	3.5	26.8	2.7	**55.1**

1 = Cohen's kappa; $\kappa = .289$, $p < .001$
2 = Cohen's kappa; $\kappa = .246$, $p < .001$
3 = Cohen's kappa; $\kappa = .349$, $p < .001$

Table 4. Longitudinal stability of stage of change for 5-a-day over 6 and 12 months

3.7 Examination of stage sequence

Stage sequence was then assessed across the 3 time points. Stable patterns were the most common for individuals in precontemplation (43.3%) and maintenance (32.5%) at baseline. The contemplation stage was very unstable with only 3.7% of respondents who began in contemplation remaining there for all three time points. Overall, more than a third (34.5%) of participants remained in the same stage over the three time points. Over the 12 month time period, about one-quarter (24.6%) of respondents progressed without relapse through the stages and one-fifth (19.3%) relapsed without progress. Precontemplators (40.4%) and contemplators (34.6%) were the most likely to progress, while maintainers (45.6%) were the most likely to relapse. Mixed patterns of change were also common with 11.7% of respondents showing the mixed inverted V pattern and 9.9 showing the mixed V pattern. Table 5 shows the patterns of stage transitions for all of the stages.

Precontemplation (n = 669)	**N**	**%**
Stable	290	43.3
Progressing	270	40.4
Relapse	N/A	
Mixed inverted V	109	16.3
Mixed V	N/A	

Contemplation (n = 81)	**N**	**%**
Stable	3	3.7
Progressing	28	34.6
Relapse	18	22.2
Mixed inverted V	19	23.5
Mixed V	13	16.0
Preparation (n = 597)	**N**	**%**
Stable	181	30.3
Progressing	115	19.3
Relapse	129	25.0
Mixed inverted V	66	11.1
Mixed V	86	14.4

Action **(n = 39)**	**N**	**%**
Stable	0	0
Progressing	9	23.1
Relapse	15	38.5
Mixed inverted V	6	15.4
Mixed V	9	23.1
Maintenance **(n = 326)**	**N**	**%**
Stable	116	35.6
Progressing	N/A	
Relapse	149	45.7
Mixed inverted V	N/A	
Mixed V	61	18.7
Overall **(n = 1712)**	**N**	**%**
Stable	590	34.5
Progressing	422	24.6
Relapse	331	19.3
Mixed inverted V	200	11.7
Mixed V	169	9.9

Note: 1. Stable – in the same stage all three time points; 2. Progressing – moving forward at least one stage with no setbacks; 3. Relapse – moving backward at least one stage without returning to the original stage; 4. Mixed inverted V pattern – participants first increased and then decreased, for example moving from contemplation to action and then back to contemplation; 5. Mixed V pattern – participants first decreased and then increased, for example moving from action to contemplation and then back to action.

Table 5. Stage sequences across 3 time points

3.8 Longitudinal prediction of stage transitions

Longitudinal prediction of stage transitions was then assessed by stage for behavior and the related psychosocial variables over the 12 month time period. All of the variables significantly predicted change across time except for subjective norms and behavior which did not predict relapse among people in action and maintenance at baseline. Self efficacy was the strongest predictor of relapse. Behavior was the strongest predictor of movement from contemplation and preparation to action or maintenance. Intention was the strongest predictor of change out of precontemplation. For transitions into stages, attitude and subjective norm predicted change from precontemplation to contemplation/preparation and perceived behavioral control predicted change from preparation to action/maintenance. All of the longitudinal predictors of stage transition are displayed in Tables 6 and 7.

		Stage at 6 months			
Baseline	PC	C/P	A/M	F value	η^2
		Behavior			
PC	1.99 (0.99)	2.19 (0.99)[1]	2.71 (1.05)[2,3]	F(2,853) = 19.02 p < .001	.045
C/P	2.53 (1.05)	2.59 (1.05)	3.15 (0.87)[2,3]	F(2,888) = 22.93 p < .001	.052
A/M	5.78 (2.00)	5.61 (1.14)	5.99 (1.35)[3]	F(2,485) = 3.45 p < .05	.014
		Attitude			
PC	8.03 (1.88)	8.57 (1.67)[1]	8.60 (1.75)[2]	F(2,858) = 9.60 p < .001	.022
C/P	8.61 (1.68)	9.01 (1.36)[1]	9.38 (0.98)[2,3]	F(2,895) = 15.92 p < .001	.036
A/M	9.25 (1.15)	9.48 (0.88)	9.61 (0.99)[2]	F(2,481) = 3.92 p < .05	.016
		Subjective Norm			
PC	6.19 (2.70)	6.76 (2.68)[1]	6.47 (2.44)	F(2,814) = 3.96 p < .05	.010
C/P	7.20 (2.47)	7.97 (2.17)[1]	8.07 (2.22)[2]	F(2,859) = 10.26 p < .001	.024
A/M	7.98 (2.04)	8.15 (1.95)	8.44 (1.82)	F(2,438) = 2.07 p = ns	----
		Perceived Behavioral Control			
PC	6.41 (2.33)	6.91 (2.28)[1]	7.12 (2.33)[2]	F(2,860) = 5.97 p < .01	.014
C/P	7.33 (2.15)	7.96 (1.82)[1]	8.43 (1.65)[2,3]	F(2,896) = 17.96 p < .001	.040
A/M	8.67 (1.68)	8.85 (1.40)	9.39 (1.01)[2,3]	F(2,481) = 14.10 p < .001	.059
		Self Efficacy			
PC	4.83 (2.88)	5.72 (2.66)[1]	5.75 (3.00)[2]	F(2,869) = 10.61 p < .001	.024
C/P	7.43 (2.21)	7.74 (2.08)	8.49 (1.85)[2,3]	F(2,907) = 14.12 p < .001	.031
A/M	8.86 (1.77)	9.04 (1.50)	9.45 (1.11)[2,3]	F(2,484) = 7.74 p < .001	.032

Note: PC = Precontemplation, C = Contemplation, P = Preparation, A = Action, M = Maintenance.
1 = C/P > PC for Tukey post-hoc test (p < .05)
2 = A/M > PC for Tukey post-hoc test (p < .05)
3 = A/M > C/P for Tukey post-hoc test (p < .05)

Table 6. Baseline behavior and mediator score by 12 month stage of change

	Stage at 1 Year				
Baseline	PC	C/P	A/M	F value	η^2
Behavior					
PC	1.98 (1.00)	2.18 (1.08)[1]	2.57 (1.06)[2,3]	F(2,719) = 9.83, p < .001	.028
C/P	2.43 (1.04)	2.68 (1.02)[1]	3.18 (0.83)[2,3]	F(2,739) = 26.1, p < .001	.066
A/M	5.70 (1.95)	5.64 (1.04)	6.00 (1.34)	F(2,389) = 2.8, p = .06	----
Attitude					
PC	7.97 (1.87)	8.63 (1.70)[1]	8.56 (1.73)[2]	F(2,721) = 11.5, p < .001	.031
C/P	8.64 (1.82)	9.02 (1.40)[1]	9.32 (1.18)[2]	F(2,746) = 9.8, p < .001	.026
A/M	9.03 (1.28)	9.63 (0.66)[1]	9.61 (1.04)[2]	F(2,384) = 10.2, p < .001	.051
Subjective Norm					
PC	6.19 (2.71)	6.74 (2.52)[1]	6.83 (2.48)	F(2,689) = 3.9, p < .05	.011
C/P	7.30 (2.51)	8.03 (2.24)[1]	7.81 (2.20)	F(2,712) = 6.6, p < .001	.018
A/M	7.97 (2.02)	8.44 (1.77)	8.33 (2.00)	F(2,353) = 1.3, p = ns	----
Perceived Behavioral Control					
PC	6.47 (2.32)	6.80 (2.32)	7.51 (1.98)[2]	F(2,722) = 5.8, p < .01	.016
C/P	7.37 (2.13)	7.95 (1.86)[1]	8.49 (1.59)[2,3]	F(2,746) = 15.7, p < .001	.041
A/M	8.30 (1.95)	9.00 (1.26)[1]	9.40 (0.95)[2,3]	F(2,385) = 19.1, p < .001	.091
Self Efficacy					
PC	4.97 (2.96)	5.75 (2.83)[1]	5.98 (2.64)[2]	F(2,733) = 7.5, p < .001	.020
C/P	7.47 (2.15)	7.72 (2.17)	8.31 (1.84)[2,3]	F(2, 757) = 7.4, p < .01	.019
A/M	8.26 (2.13)	8.98 (1.44)[1]	9.64 (0.83)[2,3]	F(2,388) = 28.1, p < .001	.127

Note: PC = Precontemplation, C = Contemplation, P = Preparation, A = Action, M = Maintenance.
1 = C/P > PC for Tukey post-hoc test (p < .05)
2 = A/M > PC for Tukey post-hoc test (p < .05)
3 = A/M > C/P for Tukey post-hoc test (p < .05)

Table 7. Baseline behavior and mediator score by 12 month stage of change

4. Discussion

This study examines the psychometric properties of the stage of change construct for fruit and vegetable consumption following three of the four research designs for testing stage models outlined by Weinstein and colleagues (1998) [33]. First, stage distribution was assessed. Over one-third of respondents were in the preparation stage of change. While this is unusual for most proactively recruited samples [25] it is consistent with other fruit and vegetable staging results [40, 27]. Next, cross-sectional comparisons of individuals in different stages showed significant differences for fruit and vegetable consumption and self efficacy as well as related constructs from the Theory of Planned Behavior including attitude, intention, subjective norm and perceived behavioral control. Attitude, self efficacy and intention all varied linearly across the stage of change. Perceived behavioral control showed a non-linear relationship with no increase between preparation and action and a large increase between action and maintenance. Attitude and subjective norms showed significant differences in the early stage with little difference between contemplation, preparation and action. Fruit and vegetable intake increased dramatically between preparation and action due to the definition of action requiring consumption of five servings of fruit and vegetables a day. These findings support the ability of stage to differentiate between related psychosocial constructs among people at different stages and support both a true stage model and the non-linear pseudostage model over the linear pseudostage model.

Stage of change was then examined longitudinally. Stage membership varied greatly over the year with less than half of respondents remaining in the same stage of change. However, the construct did show predictive validity with individuals in preparation almost three times as likely to reach action and maintenance than those in precontemplation at baseline. The preparation stage was also reached by 18% more contemplators than precontemplators at 12 months. Individuals who began the study in maintenance were also the most likely to be in maintenance at the end of the study. Participants in preparation, action and maintenance at baseline were more likely to be in action or maintenance at 12 months than those who started at precontemplation or contemplation. This supports the assumption that the stage of change for fruit and vegetable consumption are temporally ordered, with preparation the closest to action [33].

Next, we looked at patterns of change across the three time points. Stable patterns were the most common for individuals in precontemplation and maintenance. This is consistent with finding with stage of change for smoking [39]. Overall, we saw about one-third of participants remaining stable in their stage, one-quarter linearly progressing, one-fifth linearly regressing and one-fifth showing unstable patterns. With little longitudinal data available of stage transitions for fruits and vegetables it is difficult to compare or contrast these results to other studies. Although Weinstein et al (1998) recommend the examination of stage sequences even they admit that, “labeling a changing pattern of transition probabilities as gradual or abrupt is somewhat subjective, so sequence data may not be very conclusive” [33]. This appears to be the case here, with the stage transition data neither supporting nor refuting the stage model. The six month time point is probably too long for what appears to be a fairly unstable behavior.

Finally, we examined the longitudinal prediction of stage transitions. All of the variables showed some predictive validity, although the effect sizes were small. According to the TTM, cognitive strategies are important for early stage progress, while behavioral processes are important for later stage progress [21]. While this study did not examine the processes of change it does contain both experiential (attitude, subjective norm) and behavioral (perceived behavioral control) psychosocial constructs. We would therefore expect that attitude and subjective norm predict change from precontemplation to contemplation/preparation and perceived behavioral control predicts change from preparation to action/maintenance. Similar to our cross-sectional findings, the data does appear to support this assumption.

5. Conclusions

Overall, the stage of change measure for fruit and vegetable consumption is well supported by these analyses. The measure differentiates between individuals cross-sectionally, provides prediction for progress to action, and does appear to show properties relevant to an actual stage model over a linear pseudostage model. These analyses do not settle the debate between stage models and a non-linear pseudostage model, since matched and mismatched interventions are needed to examine this difference. They do, however validate the utility of the stage of change measure as an important tool for designing population interventions to increase fruit and vegetable intake. Stage of change is widely used in practice due to its utility in interventions. Stage-based expert system technology has been shown to be an important tool in reaching [41, 42]. However, Weinstein and colleagues (1998) last test of matched and mismatched interventions are still uncommon, with two small studies in physical activity and smoking not supporting improved efficacy of a stage-matched intervention [43, 44].

This study has several limitations. We did not use all of the constructs from the Transtheoretical Model including decisional balance and the processes of change. Instead several constructs from the Theory of Planned Behavior were used. Although this is not entirely consistent with the TTM, since it is a *trans*-theoretical model it is not surprising that the constructs showed similar relationships across stage to the ones postulated by Prochaska and DiClemente (1983) [45]. This approach also follows what other authors have shown, with non-Transtheoretical Model constructs varying by stage of change including attitude and expectancies [46, 47]. Also differential dropout occurred across several demographic categories. However, this appeared to be related to a younger, more mobile population rather than active refusals.

To our knowledge, this is the largest study of the fruit and vegetable staging construct to examine Weinstein and colleagues (1998) first three research designs and one of the only studies for any behavior to examine the longitudinal predictors of stage transition [33]. Several questions still need to be addressed to further improve the stage of change instrument. The most important is probably the 30 day criteria for precontemplation and the 6 month criteria for maintenance. These timeframes have been applied almost universally to every behavior that stage of change has been applied to with little rationale except historical precedence. For fruit and vegetable intake, the preparation stage is endorsed by over 40% of the population yet only 22% of these individuals were in action or maintenance after one year. This might be an area where additional examination is needed since action for an acquisition behavior is very different than from cessation of an addictive behavior. Finally, intervention research testing matched and mismatched groups is needed for fruit and vegetable intake to assess the feasibility of the stage model compared to the non-linear pseudostage model.

6. Competing interests

The authors declare that they have no competing interests.

7. Authors' contributions

JM conceptualized the paper, conducted the study, and wrote a large section of the manuscript. JB conducted the statistical analysis. CM drafted sections of the manuscript. As co-investigator, CN also participated in the conception and design of the study and provided comments on the manuscript.

8. Acknowledgements

This study was funded by the Tobacco Settlement Special Fund through a contract with the Hawaii State Department of Health. Special thanks also go to QMark Research and Polling for completing the survey interviews.

9. References

[1] National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention. Physical Activity and Good Nutrition: Essential Elements to Prevent Chronic Diseases and Obesity. [www.cdc.gov/nccdphp/overview.htm]

[2] Brousseau ME, Schaefer EJ. Diet and coronary heart disease: clinical trials. *Curr Atherosclerosis Reports* 2000; 2: 487-493.

[3] Hu, FB, Willett WC. Diet and coronary heart disease: findings from the Nurses' Healthy Study and Health Professionals Follow-up Study. *J of Nutr, Health, and Aging* 2001, 5(3): 132-128.

[4] Kromhout D, Menotti A, Kesteloot H, Sans S. Prevention of coronary heart disease by diet and lifestyle: Evidence from prospective cross-cultural, cohort, and intervention studies. *Circulation* 2002, 105; 893-898.

[5] Boden-Albala B, Sacco RL. Lifestyle factors and stroke risk: exercise, alcohol, diet, obesity, smoking, drug use, and stress. *Curr Atherosclerosis Reports* 2000, 2(2): 160-166.

[6] Renaud SC. Diet and stroke. *J of Nutr, Health, & Aging* 2001, 5(3): 167-172.

[7] Dyson PA. The role of diet and exercise in type 2 diabetes prevention. *J of Prof Nurs* 2003, 18(12): 690-692.

[8] Steyn NP, Mann J, Bennett PH, Temple N, Zimmet P, Tuomiletho J, Lindstrom J, Louheranta A. Diet, nutrition and the prevention of type 2 diabetes. *Pub Health Nutr* 2004, 7(1A): 147-165.

[9] Prentice A. Diet, nutrition, and the prevention of osteporosis. *Pub Health Nutrition* 2004, 7(1A): 227-43.

[10] Key TJ, Schatzkin A, Willett WC, Allen NE, Spencer EA, Travis RC. Diet, nutrition, and the prevention of cancer. *Pub Health Nutr* 2004, 7(1A): 187-200.

[11] Willet WC. Diet, nutrition, and avoidable cancer. *Env Health Persp* 1995, 103 (Suppl 8): 165-170.

[12] Bazzano LA, Serdula MK, Liu S. Dietary intake of fruits and vegetables and risk of cardiovascular disease. *Curr Atherosclerosis Rep* 2003, 5(6): 492-499.

[13] Feldman EB. Fruits and vegetables and the risk of stroke. *Nutrition Rev* 2001, 59(1): 24-27.

[14] Ness AR, Powles JW. Fruit and vegetables, and cardio-vascular disease: a review. *Int J of Epi* 1997, 26(1): 1-13.

[15] Van't Veer P, Jansen MC, Klerk M, Kok FJ. Fruits and vegetables in the prevention of cardiovascular disease and cancer. *Pub Health Nutr* 2000, 3(1): 103-107.

[16] World Cancer Research Fund and the American Institute for Cancer Research. *Food, Nutrition and the Prevention of Cancer: A Global Perspective.* Washington, DC: American Institute for Cancer Research; 1997.

[17] US Department of Agriculture: *USDA's Food Guide Pyramid*. Washington, DC: US Government Printing Office; 1992.

[18] US Department of Health and Human Services: *The Surgeon General's Report on Nutrition and Health.* Washington, DC: US Government Printing Office; 1988.

[19] Division of Adult and Community Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System Online Prevalence Data, 1995-2002. [www.cdc.gov/brfss]

[20] Prochaska JO, Redding CA, Evers KE. The Transtheoretical Model and Stages of Change. In *Health Behavior and Health Education*, 3rd edition. Edited by Glanz K, Rimer BK, Lewis FM. San Francisco, CA: John Wiley and Sons Inc; 2002.

[21] Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: Applications to addictive behaviors. *Am Psych* 1992, 47(9): 1102-1112.

[22] Glanz K, Patterson RE, Kristal AR, Feng Z, Linnan L, Heimendinger J, Hebert JR. Impact of work site health promotion on stages of dietary change: the Working Well Trial. *Health Ed Beh* 1998, 25(4): 448-463.

[23] Schumann A, Nigg CR, Rossi JS, Jordan PJ, Norman GJ, Garber CE, Riebe D, Benisovich SV. Construct validity of the stages of change of exercise adoption for different intensities of physical activity in four samples of differing age groups. *Am J of Health Prom* 2002, 16: 280-287.

[24] Bowen, AM, Trotter, R. HIV risk in intravenous drug users and crack cocaine smokers: predicting stage of change for condom use. *J of Consulting & Clinical Psych* 1995, 63(3): 238-248.

[25] Prochaska JO, Velicer WF, Rossi JS, Goldstein, MG, Marcus BH, Rakowski W, Fiore C, Harlow LL, Redding CA, Rosenbloom D, Rossi SR. Stages of change and decisional balance for 12 problem behaviors. *Health Psych* 1994, 13: 39-46.

[26] Brug J, Glanz K, Kok G. The relationship between self-efficacy, attitudes, intake compared to others, consumption and stages of change related to fruit and vegetables. *Am J of Health Prom* 1997, 12: 25-30.

[27] Campbell MK, Symons M, Demark-Wahnefried W, Polhamus B, Bernhardt JM, McClelland JW, Washington C. Stages of change and psychosocial correlated of fruit and vegetable consumption among rural African-American church members. *Am J of Health Prom* 1998, 12: 185-191.

[28] Ma J, Betts NM, Horacek T. Measuring stages of change for assessing readiness to increase fruit and vegetable intake among 18- to 24-year-olds. *Am J of Health Prom* 2001, 16: 88-97.

[29] Sorensen G, Stoddard A, Macario E. Social support and readiness to make dietary changes. *Health Ed & Beh* 1998, 25: 586-598.

[30] Van Duyn MAS, Heimendinger J, Russek-Cohen E, DiClemente CC, Sims LS, Subar AF, Krebs-Smith SM, Pivonka E, Kahle LL. Use of the Transtheoretical Model of Change to successfully predict fruit and vegetable consumption. *J of Nutr Ed* 1998, 30: 371-380.

[31] Van Duyn, MAS, Kristal AR, Dodd K, Campbell MK, Subar AF, Stables G, Nebling L, Glanz K. Association of awareness, intrapersonal and interpersonal factors and dietary change with fruit and vegetable consumption: a national survey. *Am J of Health Prom* 2001, 16: 69-78.

[32] Povey R, Conner M, Sparks P, James R, Shepherd R. A critical examination of the application of the Transtheoretical Model's stages of change to dietary behavior. *Health Ed Res* 1999, 14(5): 641-651.

[33] Weinstein ND, Rothman AJ, Sutton SR. Stage theories of health behavior: conceptual and methodological issues. *Health Psych* 1998, 17(3): 290-9.

[34] Waksberg, J. Sampling methods for Random Digit Dialing. *J of the Am Stat Assn* 1978, 73: 40-46.

[35] Thompson FE, Subar AF, Smith AF, Midthune D, Radimer KL, Kahle LL, Kipnis V. Fruit and vegetable assessment: performance of 2 new short instruments and a food frequency questionnaire. *JADA* 2002, 102: 1764-1772.

[36] Greene GW, Clark P, Prochaska JO, Riebe D, Nigg CR. Stage-based health promotion with the elderly. NCI, 2nd meeting of the HPRB Nutrition Behavior Grantees. Washington, DC; 2000.

[37] Montano DE, Kasprzyk D. The Theory of Reasoned Action and the Theory of Planned Behavior. In *Health Behavior and Health Education*, 3rd edition. Edited by Glanz K, Rimer BK, Lewis FM. San Francisco, CA: John Wiley and Sons Inc; 2002.

[38] SPSS (version 11.5.0). Chicago: SPSS, Inc; 2002.

[39] Prochaska JO, Velicer WF, DiClemente CC, Guadagnoli E, Rossi JS. Patterns of change: Dynamic typology applied to smoking cessation. *Multivariate Beh Res* 1991, 26(1): 83-107.

[40] Frame CJ, Green CG, Herr DG, & Taylor ML. A 2-year stage of change evaluation of dietary fat and fruit and vegetable intake behaviors of cardiac rehabilitation patients. *Am J of Health Prom* 2003, 17: 361-368.

[41] Prochaska JO, DiClemente CC, Velicer WF, Rossi JS. Standardized, individualized, interactive and personalized self-help programs for smoking cessation. *Health Psych* 1993, 12: 399-405.

[42] Redding CA, Prochaska JO, Pallonen UE, Rossi JS, Velicer WF, Rossi SR, Greene G, Meier K, Evers K, Plummer BA, Maddock JE. Transtheoretical individualized multimedia expert systems targeting adolescents' health behaviors. *Cognitive & Beh Prac*, 1999, 6: 142-151.

[43] Blissmer B, McAuley E. Testing the requirements of stages of physical activity among adults: the comparative effectiveness of stage-matched, mismatched, standard care, and control interventions. *Ann of Beh Med* 2002, 24: 181-189.

[44] Quinlan KB, McCaul KD. Matched and mismatched interventions with young adult smokers: testing a stage theory. *Health Psych* 2000, 19: 165-171.

[45] Prochaska and DiClemente. Stages and processes of self-change of smoking: Toward an integrative model of change. *J of Consulting & Clinical Psych* 1983, 51: 390-395.

[46] Jordan PJ, Nigg CR, Norman GJ, Rossi JS, Benisovich SV. Does the Transtheoretical model need an attitude adjustment? Integrating attitude with decisional balance as predictors of stage of change for exercise. *Psych of Sport & Exercise* 2002, 3: 65-83.

[47] Noar SM, Laforge RG, Maddock JE, Wood MD. Rethinking positive and negative aspects of alcohol use: Suggestions from a comparison of alcohol expectancies and decisional balance. *J on Studies on Alcohol* 2003, 64: 60-69.

3

Strategies for Cardiovascular Disease Prevention in Rural Southern African American Communities

Ralphenia D. Pace, Norma L. Dawkins and Melissa Johnson
Tuskegee University
USA

1. Introduction

Cardiovascular disease (CVD) is a commonly recognized umbrella term encompassing conditions, disease or disorders of the heart and/or blood vessels that may result in impairment of optimal cardiovascular functioning. CVD is the leading cause of morbidity and mortality in the United States, as well as in both developed and developing nations. Although the risk for CVD in the United States may vary according to and fluctuate within certain demographic, educational, behavioral and socio-economic characteristics, disparities in CVD prevention and treatment continue to ensue, particularly within minority populations. This gap in CVD disparity is further increased in regard to African American women living in rural southern communities. Demographic, socioeconomic and neighborhood characteristics are suggested to converge within rural African American communities to additively influence CVD risk.

Efforts to prevent CVD have often involved the use of diet and nutrition, nutrition education, physical activity modifications and behavioral-based strategies. The higher rate of disparities in CVD prevalence, mortality, preventive and treatment services in rural regions warrants an aggressive approach in addressing this issue to improve public health. The use of prevention strategies, exploiting specific aspects of the above mentioned strategies may prove useful in mitigating CVD risk disparities among African Americans living in rural southern communities. This chapter will investigate CVD risk, prevention and treatment, with an emphasis on African Americans living in rural southern communities. Additionally, several strategies employing an integrative multi-disciplinary approach to preventing CVD in rural southern African American communities will be provided.

2. Definitions and classification of CVD

CVD is an umbrella term encompassing conditions such as: high blood pressure (HBP); coronary heart disease, including myocardial infarction and angina pectoris; heart failure, stroke and congenital cardiovascular defects (Lloyd-Jones et al., 2009). Classifications of CVD include, but are not limited to: atherosclerosis, cardiomyopathy, cerebrovascular disease (stroke), congenital heart disease, coronary heart disease, hypertension, heart failure, and transient ischemic attack. Ideally to achieve and maintain cardiovascular health several anthropometrical (e.g. body mass index), behavioral (e.g. physical activity), dietary (e.g. healthy diet score) and clinical (e.g. total cholesterol, blood pressure and fasting plasma glucose) objectives should be met (Lloyd-Jones et al., 2010).

2.1 CVD prevalence in the United States

Cardiovascular disease (CVD) has been recognized as a public health problem for nearly a century in the United States, whereas in previous years contagious diseases such as typhoid disease, smallpox, diphtheria, rheumatic fever and tuberculosis posed a particular threat to public health (Griswold, 1927). Over one-third of adults living in the United States over the age of 20 years have high blood pressure, one of the most common types of CVD (Roger et al., 2011). Among these individuals the highest prevalence of hypertension is observed among African Americans- particularly those living in the rural South (Danaei et al., 2010). Further, these individuals also exhibit greater disparities in smoking, elevated blood glucose and adiposity. It is predicted that nearly half of individuals living in the United States will exhibit some form of CVD by 2030 (Heidenreich et al., 2011). The financial burden associated with the costs of CVD treatment, morbidity and mortality are projected to reach over $800 billion during this time.

An estimated 1 in 3 American adults are predicted to possess one or more forms of CVD, the most common of which is hypertension or high blood pressure (HBP) (Lloyd-Jones et al., 2009). Although the prevalence of heart disease is slightly less in African Americans compared to Whites (10.2% vs. 11.4%), a significantly greater number of African Americans have high blood pressure (31.7% in African Americans vs. 22.2% in Whites). Individuals living in the southern region of the United States referred to as the "stroke belt" have the highest rates of high blood pressure (Hajjar & Kotchen, 2002). Further, high blood pressure increases the risk for cerebrovascular disease or stroke, which is the third leading cause of death among Americans (Lloyd-Jones et al., 2010). Although stroke-related deaths have declined in recent years, individuals living in the southeastern region of the United States exhibit the highest rates of stroke and related hospitalizations.

2.2 CVD prevalence in rural southern African American communities

As mentioned previously, individuals living in the stroke belt, a cluster of communities in Arkansas, Louisiana, Mississippi, Alabama, Georgia, South Carolina, North Carolina, Virginia, Tennessee, Kentucky and Indiana, exhibit stroke death rates higher than the national average. In comparison to other regions of the United States, CVD prevalence is elevated among individuals living in rural southern communities, particularly African American women (Taylor et al., 2002). The highest rates of CVD mortality are often observed among those who are poor and live in rural regions (Cooper et al., 2000).

2.3.1 Demographic and socioeconomic characteristics of rural communities

Rural communities are generally composed of individuals who are older, have lower educational attainment and lower socioeconomic status (Johnson, 2006). Persistent sluggish economies and lower income brackets have contributed to the continual presence of poverty across generations observed in rural America (Brown & Warner, 1991). Although a vast majority of non-Hispanic Whites occupy rural communities, rural communities in the southeastern United States are heavily populated by African Americans.

Although younger residents (21 to 35 years of age) of rural communities demonstrate greater knowledge regarding the general relationship between diet and CVD risk, these individuals were also found to express less immediate concern in developing adverse health outcomes (Pace et al., 2008). Nevertheless with knowledge of the relationship between a specific dietary component (i.e. fat) and CVD risk, came dietary behaviors that may reduce the risks associated with CVD (i.e. consumption of low-fat dairy products). Among 18 to 26 year old African Americans living in rural communities, females and greater educational attainment were associated with greater CVD risk knowledge (Winham & Jones, 2011). Although rural African American women may perceive themselves at risk for hypertension, barriers such as income, lack of insurance, medical expenses, neighborhood environment and family support/characteristics may inhibit them from actively seeking CVD prevention and/or treatment measures (Ford et al., 2009).

2.3.2 Rural African American neighborhood/community resources

In comparison to urban communities, rural communities have limited access to fruits and vegetables (Hosler et al., 2008). Research indicates that neighborhood characteristics influence the affordability of fresh fruit and vegetables, with African American rural residents paying more for these items (Dunn et al., 2011). Perceptions regarding self-efficacy, the neighborhood (i.e. community) and home (i.e. consumer) nutrition environment and family support among rural Georgia residents indicate positive associations between these variables and healthy dietary behaviors (Hermstad et al., 2010). Components of the neighborhood nutrition environment include food accessibility as influenced by the number, type and location of grocery/convenience stores and restaurants. Conversely, the home nutrition environment includes the presence, cost and quality of healthy food items in the neighborhood, which influence consumer nutrition behaviors such as shopping at supermarkets versus convenience stores and dining out at a sit-down versus a fast-food establishment. The home nutrition environment and consumer nutrition behavior were positively associated with dietary behaviors that facilitate dietary fat intake among rural women.

Senior citizens living in rural communities are at risk for inadequate fruit and vegetable intake as well. Older women reported consuming more servings of fruit and vegetables, although both men and women reported consuming comparable servings of fruit and vegetables in rural Texas (Sharkey et al., 2010). Likewise, fruit and vegetable consumption among these individuals was influenced by supermarket location, produce variety and produce quality. Consequently, the diet quality of rural residents is compromised. Data collected on the diet quality of older adults living in the rural southern United States revealed that less than 2% met the dietary guidelines (Savoca et al., 2009). Although overall diet quality was inadequate, dietary intakes of dark green and orange vegetables were sufficient. In comparison to American Indians and in some instances non-Hispanic whites, rural African Americans reported consuming greater quantities of total and whole fruit, grains and meat.

Sanderson et al (2003) found that nearly two-thirds of African American women living in rural Greene, Lowndes and Wilcox counties of Alabama were either “insufficiently inactive” (46%) or “inactive” (15%). Contributing to their current level of physical activity or inactivity were participation in regular religious services, observing others engaging in physical activity (i.e. exercise), interacting with others engaging in physical activity and more positive feelings and perceptions regarding participation in physical activity. In addition, an immense sense of community related to neighborhood characteristics such as quality, safety and neighbor support contributed to more personal positive feelings and perceptions regarding physical activity. Rural women who viewed their neighborhood as safe and observed the presence of sidewalks were more likely to engage in regular physical activity (Wilcox et al., 2003). Glover et al (2011) found rural South Carolina female children to be significantly more likely to engage in physical activity than their male counterparts. Among these children less than half (42%) indicated that they consume fewer than the recommended daily servings of fruit and vegetables.

2.4 Racial/ethnic disparities in CVD

Racial and socioeconomic distinctions between CVD risk among women living in the rural southern region of the United States have been noted by researchers (Appel et al., 2002). Women with lower education levels exhibited the greatest risk for CVD, as augmented by inadequate physical activity, smoking, elevated cholesterol levels and a family history of CVD. In comparison to White women, African American women living in rural communities had significantly lower education and income levels. Among these women the prevalence of hypertension and diabetes was greater among African American women. Although deaths from CVD have declined in individuals living in rural regions, those in the southern and Appalachian regions of the United States are still at an amplified risk for premature death from CVD (Pearson and Lewis, 1998; Barnett & Halverson, 2000; Barnett et al., 2000). National rates of death in the United States from heart disease were the second highest in southern rural counties; among men the highest percentage of deaths related to CVD were observed in these counties (Eberhardt et al., 2001).

Disparities in access to medical treatment facilities in rural areas are believed to contribute to the increased risk of premature death from CVD among rural residents. In comparison to urban communities, rural communities are more likely to have inadequate access to and quality of health care services (Reschovsky & Staiti, 2005). Minorities (i.e. African Americans and Hispanic Americans) living in rural regions have disproportionately limited access to health care compared to their White counterparts (Mueller et al., 1999). African American women living in rural areas are particularly vulnerable (Cort et al., 2001). Significant disparities in access to facilities that provide treatment for acute cases of CVD (i.e. stroke) have been observed when comparing urban to rural communities, with significantly fewer rural communities having access to acute care facilities (Khan et al., 2011).

Disparities in stroke prevalence have also been observed with African Americans being more likely to report having a stroke, compared to other ethnic groups; African Americans were also more likely to report being hypertensive (McGruder et al., 2004; Lloyd-Jones, 2009). Among these individuals, black non-Hispanic females with less than a high school education, living above or equal to the poverty line with an annual income <$20,000 and unemployed were more likely to have a stroke. Racial/ethnic disparities in CVD risk in regard to the presence of other co-morbidities such as diabetes, insulin resistance and hypertension have also been observed (Brown, 2006).

Zuniga et al (2003) suggest that disparities in CVD among rural Americans may be augmented due to certain behavioral characteristics and attitudes. Lower educational attainment, socioeconomic status and standards of living among individuals in rural communities may contribute to behaviors that increase CVD risk; poor dietary patterns, smoking, physical inactivity and the failure to manage other co-morbidities such as hypertension and diabetes contribute to this increased risk (Cooper et al., 2000).

2.5 Classical and novel assessments of CVD risk

CVD clinical risk assessment has customarily been assessed based on an individual's lipid profile. Generally, the risk for CVD is amplified due to the risk for and presence of other co-morbidities. In addition to the more common physiological manifestations of CVD risk, novel assessments of CVD include the presence of more recently recognized proteins and cytokines that influence CVD risk. It has been suggested that acute phase proteins such as C - reactive protein (CRP), fibrinogen, homocysteine and lipoprotein (a) may be useful as novel assessment parameters of CVD risk (Hackman & Anand, 2003).

Conventional risk factors for CVD risk assessment include biological (e.g. genetic predisposition, physiological (e.g. inflammatory disorders, hypertension, hyperlipidemia, diabetes), behavioral (e.g. cigarette smoking, physical inactivity) (Khot et al., 2003), dietary (e.g. inadequate fiber and antioxidant nutrients, high fat composition) and demographic (e.g. age, race, sex, education, income) (Figure 1). Indirect CVD risk may be assessed based on the presence of other co-morbidities as certain disease states increase the risk for developing CVD. For example, persistent elevations in blood glucose as seen in diabetes mellitus may increase the risk for hypertension. As indicated earlier in regard to racial/ethnic disparities in CVD risk, hypertension typically afflicts African Americans earlier in life and with greater severity.

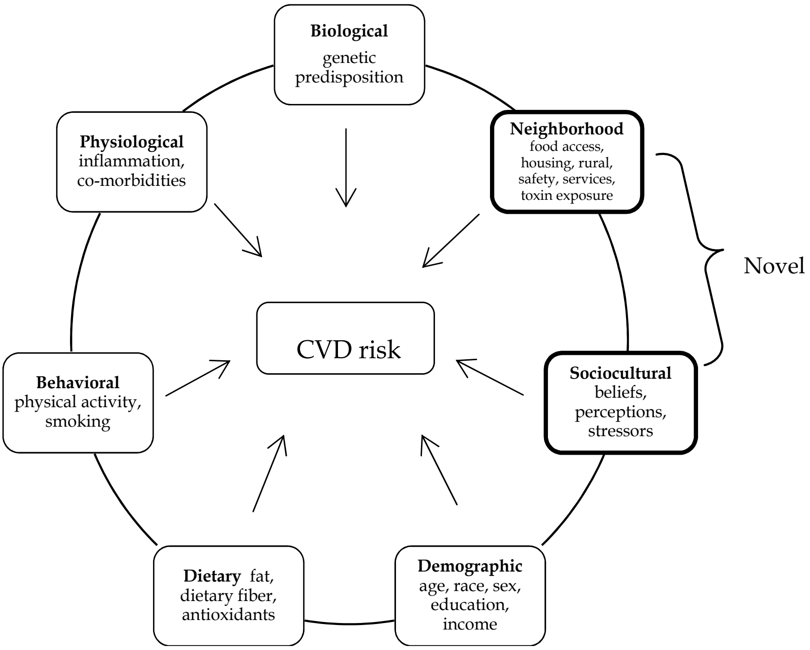


Fig. 1. Conventional and novel influencers of CVD risk.

The assessment of CVD risk in African Americans is elusive in that generational non-genetic trajectories not commonly observed among other subgroups ensue in the African American population. Because African Americans typically exhibit lipid profiles within the "normal" range, an assessment of CVD risk (or other co-morbidities) based solely on lipid profile may not be sufficient to determine the level of true risk (Sumner et al., 2005a). While African Americans may display normal triglyceride levels even in the presence of insulin resistance, these individuals may have increased lipoprotein activity, which serves to attenuate increases in postprandial triglyceride levels (Sumner et al., 2005b). Although triglyceride levels may be used as a clinical diagnostic tool for insulin resistance, the demonstration of normal triglyceride levels among African Americans with insulin resistance (Sumner & Cowie, 2008), suggest that the identification of this disorder, as well as CVD, based solely on these variables may be insufficient. The integration of several specific biomarkers for disease risk as well as socioeconomic and demographic characteristics that influence disease risk may prove useful in the assessment of CVD risk among African Americans.

Research findings suggest that socioeconomic characteristics may influence and/or mediate the manifestation of certain physiological processes, which may increase the risk of adverse cardiovascular outcomes (Muennig et al., 2007; Aiello & Kaplan, 2009). Further, chronic socioeconomic conditions may transcend generations and result in persistent inflammatory and immunologic responses that increase CVD risk. Lower socioeconomic characteristics during childhood have been associated with CVD in adulthood (Galobardes et al., 2006). Lower socioeconomic status has been found to be positively associated with inflammatory biomarkers such as fibrinogen and C - reactive protein (Wilson et al., 1993; Tabassum et al., 2008). Predictors of socioeconomic status and certain biomarkers related to CVD risk include employment/occupational status, educational attainment and income (Yarnell et al., 2005; Muennig et al., 2007; Ranjit et al., 2007; Rosvall et al., 2007). Of the predictors of socioeconomic status (i.e. education, income and occupation), education was significantly associated with decreased blood pressure, decreased total cholesterol and increased high-density lipoprotein cholesterol (HDL-C) (Winkleby et al., 1992). Albert et al (2006) found education and income to function as novel assessments of CVD risk, with education and income being inversely associated with CVD risk.

3. Global perspective of CVD

Cardiovascular disease (CVD) is caused by disorders associated with the heart and blood vessels, which include heart attack (coronary heart disease), stroke, hypertension, peripheral artery disease, rheumatic heart disease, congenital heart disease, and heart failure (WHO, 2011). According to the World Health Organization (WHO) report, 17.3 million people died from CVD in 2008. Of those deaths, 7.3 million resulted from coronary heart disease and 6.2 million from stroke. Low and middle income regions are disproportionally affected with over 80% of the deaths occurring in these locations. It is projected by 2030 that approximately 23.6 million people will die from CVD, which will remain the leading cause of death (WHO, 2011). Despite the declines in CVD over the past few decades, it still remains the leading cause of death in the United States.

In 1992, more than 816,000 Americans died from some form of CVD, compared to 631,636 in 2006. In this same year heart disease was the major cause of death among various ethnic groups as well as for the general population. African Americans had the highest death rates compared to Asian and Pacific Islanders, Native Americans or Alaskan natives and Hispanics. In general heart disease was highest in Mississippi, a Black Belt state characterized by a population density of African Americans (CDC, 2010). It was estimated that in 2010 heart disease would cost the United States 316.4 billion dollars. This cost includes: healthcare services, medications, and lost productivity (CDC, 2010). Many Black Belt states exhibit higher prevalence of CVD, hypertension or high blood pressure (HBP) and obesity that exceed the national average (Table 1). It is estimated that 9 out of ten individuals with heart disease have at least one risk factor. Modifiable risk factors high blood pressure, high cholesterol, diabetes, smoking, overweight and obesity, poor diet, physical inactivity and alcohol use, are associated with lifestyle behaviors (Fakiri et al., 2006). Awareness of modifiable risk factors can lead to positive improvement in the health of individuals.

3.1 Stroke, hypertension and obesity

African Americans are disproportionally affected by CVD, and are more likely to have two or more risk factors (Gillum, 2001). Furthermore research studies have shown that African-American children have an increased relative stroke risk of 2.12 compared to Whites (Fullerton 2003). Also death rates for stroke was 48.1 for White males compared with 73.9 for African-American males; the disparities followed similar trends for White females, with a rate of 47.4 compared to 64.9 for African-American females per 100,000 (Fullerton, 2003). In general African Americans have been shown to be at risk 70% more than whites for stroke hospitalization (Kennedy, 2002). The severity, disability and mortality from stroke for African Americans are greater compared to Whites.

African Americans develop hypertension at an earlier age than Whites and Mexican Americans (CDC, 2010; Ong et al., 2007). Among African Americans more females (44.1%) than males (42.2%) tend to develop hypertension. The percent affected by hypertension is higher for African-Americans males and females 42.2% and 44.1% compared to the general population 31.8 and 30.3, respectively. The prevalence of hypertension in African Americans in the United States is among the highest in the world (Hertz, 2005). Hypertension is a major risk factor for heart disease, stroke, congested heart failure and kidney failure (CDC, 2010). Within the African American community, rates of hypertension vary substantially. Those with the highest rates are more likely to be middle aged or older, less educated, overweight or obese, physically inactive, and have diabetes mellitus (Collins, 2002).

In the past four decades the prevalence of obesity among U. S. adults increased from 13 to 32%. Presently, 66% of adults are overweight or obese; 16% of children and adolescents are obese and 34% are overweight. Overweight adolescents have a 70% chance of becoming overweight or obese adults and this number increases to 80% if one or both parents are obese (Kaufman, 20007). It is projected that by 2015, 75% of adults will be overweight and 41% will be obese (Wang and Beydoun, 2007). According to the Centers for Disease Control and Prevention none of the states in the United States of America met the Healthy People 2010 's goal to reduce obesity prevalence to 15%. Furthermore, the number of states with an obesity prevalence of 30% or more increased to 12 states in 2010 (CDC, 2011). In general the states where African Americans are present in the highest concentrations, They tend to have higher rates of obesity. Additionally, higher rates of obesity are found among groups with low educational and income levels, racial and ethnic minorities, rural and high poverty areas. Obesity is a risk factor for CVD and other chronic diseases and disproportionally affects African Americans of all ages (Odgen et al., 2006). The rate of obesity is higher in rural areas, due in part to poor nutrition, physical inactivity, and low educational levels (Jackson, Doescher, Jerant & Hart, 2005). In Table 1 below, the prevalence of CVD is illustrated in the Black Belt States.

	CVD[1](%)	HBP (%)	Obesity (%)
Alabama	4.0	34.0	32.2
Arkansas	5.2	32.5	30.1
Georgia	3.8	31.7	29.6
Florida	4.1	28.3	25.6
Louisiana	5.0	34.3	31.0
Mississippi	3.9	36.2	34.0
North Carolina	4.2	30.4	27.8
South Carolina	4.1	31.0	31.5
Tennessee	4.0	30.6	30.8
National	**4.5**	**29.7**	**33.8**

Table 1. CVD, HBP and obesity prevalence[1] in adults across the Black Belt States.

3.2 Rural communities defined

Because there are different characterizations of rural areas, there is no single preferred definition that suits all policy requirements. The US Census Bureau defines an urbanized area asan area that includes a central city and the surrounding densely settled territory that together have a population of 50,000 or more and a population density generally exceeding 1,000 people per square mile. The Office of Management Bureau designates areas as metro; Economic Research Service/United States Department of Agriculture (USDA) uses rural-urban continuum codes to distinguish metro counties by size and non-metro counties by their degree of urbanization or proximity to urban areas. USDA uses codes zero to three as metro; four to nine as non-metro. An urban population with a designation of four has 20,000 inhabitants or more adjacent to a metro area and a code of nine signifies completely rural or urban with a population of fewer than 2,500 not adjacent to a metro area (USDA National Agricultural Library, 2008).

3.2.1 Rural African American communities

For the purpose of this text rural areas/communities are smaller towns or cities with low population density; where most of the land is devoted to agriculture. This definition better describes the communities inhibited by a vast majority of African Americans. The Black Belt region is characterized by rural socioeconomic decline, inadequate programs, acute problem of poverty, poor health, substandard housing and underemployment. This region includes the southern most states with a high percentage of African Americans/Blacks (Webster & Bowman, 2008; Wimberly & Morris 1997). A large segment (54.8%) of the African American

---

[1] Coronary Heart Disease Prevalence

population resides in this region compared to 17.5%, 18.8%, 8.7% in the northeast, Midwest, and West, respectively (US Census Bureau 2000). For example, in the state of Alabama there are 17 Black Belt counties which include Barbour, Bullock, Butler, Choctaw, Dallas, Greene, Hale, Lowndes, Macon, Marengo, Montgomery, Perry, Pike, Russell, Sumter, and Wilcox (Figure 1). The total population in this region is close to 600,000 or about 13 percent of the 4.5 million in the State**.** The percent of poverty ranges from 26.8% to 40% within the Black belt counties (Federal Statistics, 2004). The per capita income for the State of Alabama is estimated at $33,945, but is much lower in selected Black Belt counties with median incomes ranging from $24,969 to $30,370. Furthermore, median incomes for non-Black Belt counties in Alabama ranged from $41,770 to $64,371. When compared with other Alabama counties, Black Belt counties in Alabama have disproportionate greater rates of heart disease, cancer, hypertension and diabetes. CVD mortality rates are higher in Black Belt counties compared to non-Black Belt counties in Alabama (Table 1). Similar trends exist in the Delta region. In general CVD mortality rates are higher, the rate in Alabama is 235, Mississippi is 267.6 and South Carolina is 306 (from 2004 report)compared to the national average of190 per 100,000.

🟧 **Traditional Counties of the Alabama Black Belt**

Source: Center for Business and Economic Research,
The University of Alabama

Fig. 2. Alabama black belt counties.

Black Belt States	Income
Alabama	33,945
Arkansas	33,150
Georgia	35,490
Florida	39,272
Louisiana	38,446
Mississippi	31,186
North Carolina	35,638
South Carolina	33,163
Tennessee	35,307
Virginia	44,762
National	**40,584**

Table 2. Per capita personal income (2010) of households in the Black Belt States.

County	Health Ranking*	Median Income	High School Graduation Rate	CVD Mortality Rate**
Black Belt				
Macon	61	27,011	57.6	373
Bullock	67	24,969	65.9	323
Barbour	31	30,370	53.5	295
Lowndes	62	28,530	48.3	400
Non-Black Belt				
Lee	2	41,770	69.2	178
Shelby	1	64,371	76.2	145
Jefferson	29	43,279	60.8	265
Alabama				235
National	-	50,221	84.6	191

*County Health Ranking 1= best performance and 67 = worst performance.

**CVD mortality rate per 100,000 Taken from "*Selected Indicators of Health Status in Alabama", AL Rural Health Association and Alabama Department of Public Health, 2007*

Table 3. Comparison of Four Targeted Black Belt Counties and Non Black Belt Counties in Alabama on Selected Health Measures

Americans living in rural areas are disadvantaged as it relates to healthcare. These rural residents have to travel long distances to reach a healthcare facility, have less access to specialized care and are less likely to receive preventive care (Larson and Fleishman, 2003). Data showed that rates were higher (176.3) among rural African Americans admitted for uncontrolled diabetes without complications compared to metropolitan Whites and Blacks - 13.8 vs. 76.7, respectively (AHRQ, 2005). In general pre-mature mortality is greater among rural residents than among urban/metropolitan or suburban. The age-adjusted death rate among individuals aged 1-24 who lived in rural counties was 31% higher than those living in urban counties and 65% higher than individuals living in suburban counties. Similar trends exist for more mature adults living in rural counties when compared to their urban counterparts. Death rates from CVD and cancer are higher in rural areas within certain regions. The mortality rate from heart disease was highest in the South (Eberhardt et al., 2001); furthermore, it was 25% higher than the rate among Southern suburban residents. The gap becomes wider for African American women where a 10-fold difference was observed; and the highest rates were among residents who lived in rural areas in the Mississippi Delta region (Taylor, Hughes & Garrison, 2002). Stroke is also higher among rural African Americans (Gillum, 1997). In Mississippi 80% or more of the counties have no physicians who specialize in CVD. Furthermore there is a lack of medical care resources for coronary care unit beds and cardiac rehabilitation units as well as limited action for intervention and treatment. Additionally, an analysis of services to Medicare beneficiaries revealed that the level of cardiology services for rural Medicare recipients was 40% lower than the urban beneficiaries as a result of the lower number of doctor service per beneficiary (15%) (Taylor, 2002). Furthermore, there is an increasing body of research that supports the presence of physical bias on race/ethnicity, economic status and in some cases gender (Fincher et al., 2004).

3.3 CVD interventions

The Department of Health and Human Service in 1990 established the *Healthy 2000* National Health Promotion and disease Prevention objectives, a strategy for improving the health of Americans. In 2000, the 2010 objectives were launched, as "a comprehensive nationwide promotion and disease prevention agenda. There were 467 objectives designed to serve as a frame work for improving the health of all people in the United States. In December of 2010 the *Healthy People* 2020 was delivered which continues to build and expand goals and objectives established two decades earlier (CDC, 2011). The objectives relating to reducing health disparities in African Americans was partially met; for example diabetes prevalence, diabetes-related deaths and lower extremity amputations. However, fetal alcohol syndrome increases and the gap widened (CDC, 2002). Disparities among African Americans generally continue in all aspects of their lives. The Healthy People indicated that community-based intervention is an important method for achieving health objectives. Community-based partnerships have effectively changed health related issues such as establishing requirements for smoke free schools and labeling of heart healthy foods (Brownson, 1996). Large community based interventions have addressed both individual and community wide changes with emphasis on CVD prevention. These large interventions include: the North Karelia project, Finland, Stanford Five-City Project, and Pawtucket Heart health Program. The projects in the US reported favorable results except for the Minnesota Heart Study, where significant progress in reduction of risk factors were not realized. Unfortunately none of these large scale cardiovascular studies were focused on African Americans or rural populations (Brownson et al., 1996).

Interventions targeting African Americans are sparse, however smaller intervention projects focused on the rural African American population have demonstrated promising results. Significant reductions in body weight, total cholesterol, diastolic and systolic blood pressure among African Americans who participated in a Nutrition Education Program in rural Alabama were observed (Qian et al., 2005). Increased folate intake among rural African-American men at risk for CVD was also reported after 12 weeks of nutrition education.

Furthermore, fruits and vegetable consumption increased among rural and inner city participants (O'Loughlin, 1999; Brownson et al., 1996; Joshu et al., 2003; Resnicow et al., 2004). In the Bootheel Heart Project, an intervention focused on improving lifestyle factors related to CVD among at-risk African Americans, noted improvements for each of the five risk factors associated with CVD (i.e. leisure physical activity, smoking status, consumption of five fruits and vegetables daily, weight status and cholesterol levels). The Bootheel project was a physical activity intervention that focused on African Americans at risk for CVD as a result of physical inactivity and other lifestyle factors.

Increasing physical activity in the rural African American population is a major factor in partially addressing CVD health disparities issues. For physical activity intervention strategies to be effective, barriers and limitations must be identified and addressed. The most often cited reason for physical inactivity is the lack of facilities. In the rural areas, there are no designated walking trails or sidewalks. Safety is sometimes cited (Carter, 2009 unpublished data). Eyler et al (2003) reported positive outcome in physical activity levels among African American women involved in a cardiovascular Health Network project. In one study although an increased use of the walking trail was reported a significant change in walking rates was not evident (Brownson et al., 2004) A Faith-based Institution engaged in physical activity programs has seen a 54% increase in churches implementing such programs. Physical activity is an important factor for CVD prevention (Wilcox et al., 2006). It will require creative approaches to engage rural African Americans in physical activity and other life changing habits that are sustainable to achieve desired goals in CVD disparity reduction

3.4 CVD prevention among African Americans

The reasons for disparities observed among the African American population are complex and often interrelated. They are associated with low socioeconomic status, inequalities in work, income, education, limited access to health care and overall standard of living. African Americans living in the South and specifically the Black Belt/ Lower Mississippi Delta region have the highest CVD rates in the nation. Much effort has focused on pharmacologic management of CVD, although these treatments have proven benefit, they are costly and may have side effects and may require additional medical intervention. As noted earlier there are many rural areas without physicians; and access to cardiac physicians is a limitation for this segment of the population.

CVD prevention programs for African Americans especially in the rural areas should be established upon the community-based participatory research model. Community-based participatory model is a constructive research paradigm use to promote active community involvement that shapes the research and intervention strategies as well as implementation of the study. It is an interactive process, incorporating, research, reflection, and dynamic action involving individuals from all levels (community leaders, participants). There are several facts that drive the need for culturally sensitive prevention strategies that are region and community specific. The approach should elucidate information on nutritional and physical activity behaviors, group support, self efficacy, socialization patterns, willingness to participate in long term lifestyle modification etc. A holistic approach is warranted which will include changes at the systems levels, to include policy changes as it related to federal appropriation and implementation of these policies. It is important to increase awareness and engage African Americans in culture specific activities, broaden their perceptions and bring light to the disparities and action that will reduce these anomalies.

4. Disease prevention strategies for rural African Americans

African Americans living in rural America would benefit most from disease prevention program strategies within their communities since in comparison to other ethnic groups they are disproportionately affected by CVD, particularly hypertension, when compared to their White counterparts. They are also a group with less access to overall health care. Rural communities in Black Belt states (AL, FL, GA, LA, NC, SC, MS, TN and AR) throughout the US generally have higher ratesof CVD per 100, 000, often exceeding the national rate. These statistics obviously show the need for intervention within these communities.

Place of geographical residence has also been implicated as a factor in determining health status. In an evaluation among rural, urban and suburban residence, individuals living in the most urban and most rural areas were the most disadvantaged relative to health measures (Eberhardt & Pamuk, 2004). In a study conducted by Mainous III et al (2004) in Charleston, South Carolina, a Black Belt state, on race, rural residence and control of hypertension, the results strongly suggested that among patients with diagnosed hypertension, 11% of rural Whites, 13% of urban Whites, 20% of urban African Americans, and 23% of rural African Americans had diastolic blood pressures greater than 90 mmHg ($P<0.01$).

In addition to risks for disparities in health, living in rural communities creates other disadvantages such as availability and accessibility to healthier food selections. Food availability and accessibility factors were evaluated in two contrasting cities in Tuskegee, AL (located in Macon County, a Black Belt County) and Auburn, AL. Thirty retail outlets were evaluated for the availability of selected foods in Macon County. More healthy food selections such as frozen, low-sodium or dark-green, yellow vegetables, low-fat milk or yogurt, low-sodium and low-fat cheese were often unavailable in convenience stores; none of the supermarkets in the same location stocked low-sodium vegetables (Bovell-Benjamin et al., 2008). Similar findings on availability and accessibility were found in a South Los Angeles restaurant relative to African Americans having healthy food options, both in food selections and in food preparation. Restaurants in economically disadvantaged poorer rural communities heavily promoted unhealthy food options to residents compared to residents living in more affluent areas (Lewis et al., 2005).

4.1 Diet and nutrition prevention strategies

Many disease prevention strategies for cardiovascular exist with varied results. The main stream intervention strategies emphasize the importance of nutrition education using several social theories to enhance the change process. Diet has also been implicated as a tool to reduce or prevent selected types of CVD. Consumption of diets rich in fruits and vegetables, containing significant amounts of antioxidants, high dietary fiber, low saturated and trans fat and a balance in other essential nutrients are recommended to reduce the risk for CVD. Sweet potato greens and purslane, two novel foods, and several traditional foods included in the African-American diet such as butter beans, purple hull peas, muscadine grapes, collards, butter peas, figs, okra, mustard greens, green onions, rutabagas, and eggplant are examples of foods that contain significant levels of antioxidants (Huang et al., 2007a, 2007b, 2008, 2009). In an article by Johnson and Pace (2010),the nutritional characteristics of sweetpotato (*Ipomoea batatas*) leaves were reviewed in terms of health promotion and disease prevention. The supply an abunadnce of vitamins, minerals, antioxidants, dietary fiber and essential fatty acids.

Additionally, the bioactive compounds contained in this vegetable occupy a role in promoting health by improving immune function, suppressing cancer cell growth and reducing oxidative stress and free radical damage, which are associated with the development of cardiovascular and other chronic diseases. Currently sweet potato leaves are consumed in Asian and African countries; limited consumption occurs in the United States. Additionally, the articleexamined the nutritional characteristics and bioactive compounds within sweetpotato leaves that contribute to health promotion and disease prevention. Dawkins et al (2010) found that purslane (*Portulaca oleracea*) contained relatively high amounts of omega-3 fatty acids, protein and dieary fiber and low amounts of total fat. High dietary fiber and low dietary fat are synergistic in the reduction and prevention of cardiovascular disease. However, for all of these foods,theirhealth benefits are not well known nor have their nutrient value been emphasized as excellent foods for consumptionto prevent diseases in the African American community.

4.2 Other CVD prevention strategies

After one year of participation in one of three churched-based intervention strategies, a standard behavioural group intervention, the standard intervention supplemented with spiritual strategies, or self-help strategies, 529 African American women from 16 different churches who participated in the intervention exhibited significant improvements in body weight, waist circumference, systolic blood pressure, total dietary fat and sodium intake. The self help group did not show improvement. The improvements in the intervention group suggest that theyreduced their CVD risk profiles one year after the initiation of the program. Further suggested from this research is that church-based programs can significantly benefit the cardiovascular health of African American women (Yanek et al., 2001)

4.2.1 Rural CVD prevention strategies

Nutrition, Health and Physical Activity Fairs (NHPAFs) are often used to provide information to the community to create awareness, education, and action to reduce CVD risk factors in individuals attending the fair. The NHPAFs, hosted in Macon County, AL in 2008, 2009, 2010 and 2011 were community nutrition outreach activities with residents from other neighboring counties participating. Other counties included: Bullock, Montgomery, Chambers, and Lee. These nutrition fairs were sponspored by the Department of Food and Nutritional Sciences, Tuskegee University, Tuskegee, AL.

At the NHPAFs, participants most unique experience allowed them to have onsite nutrition assessments, clinical evaluations measurements of blood pressure, glucose and cholesterol, percent body fat, waist and hip circumference, hydration level, weight, vision, hearing, as well as foot health and care and breast self-examination. Consultation with nutritionists, physicians, nurse practitioners and diabetic educators were also available for interpretation of clinical measurements. Additionally, participants were exposed to gardening and herb displays, NUTRIFOODs (sweet potato leaves, purslane, butter beans, purple hull peas, muscadine grapes, collards, butter peas, figs, okra, mustard greens, green onions, rutabagas, and eggplant) tasting.

Data collected at these fairs showed that individuals with scores higher than 10 (determined from a 25-question instrument where lower scores are better) were good candidates for nutrition counseling. The mean score was 18 among those 14 to 80 years old. Other outcome measures showed significant positive associations ($P < 0.05$) between weight and blood glucose, and body mass index (BMI) and systolic blood pressure; BMI was positively and significantly ($P < 0.01$) associated with blood glucose, diastolic blood pressure, and percent body fat. High nutrition scores correlated with increased risk for chronic disease (NHPAF, 2007).

Often in small rural communities with limited health resources, nutrition and health fairs can be used to assess the health status of residents. Twenty-two men, 45 women and six children participated in a one-day health fair sponsored by a school of nursing, University of Alabama. The results indicated that participants need assistance with health promotion in several areas: weight loss/obesity, blood sugar control, lowering cholesterol levels, vision and hearing follow-up, etc. (Lyons et al., 2001).

4.3 Effective social models and compliance tools for disease prevention in African Americans

The traditional social models, e.g. Social Cognitive Theory, Health Belief Model, and the Transtheoretical Model for Change etc. are commonly used to promote sustained behavioral change. Social Cognitive Theory variables (e.g. social support, self-efficacy, outcome expectations andself-regulation) are often used to reinforce desired behaviors. The question of whether all behavior is the same in African Americans compared to Whites living in rural or urban American communities remains unanswered. In a university-neighborhood health care center intervention to promote the Dietary Approaches to Stop Hypertension (DASH) diet study, 82 low-income African American adults with poorly controlled blood pressure participated (12 to 15 participants per group) in the study for eight weeks for one to two hours weekly. The intervention followed constructs of Social Cognitive Theory and featured dinners developed using the DASH diet plan. Following the dietary intervention, blood pressure was significantly decreased ($P < 0.05$) among participants who did not miss more than 2 of 8 sessions (Rankins et al., 2005).

Through the 12-week Tuskegee University Nutrition Outreach Program (TUNOP), the effects of a church-based nutrition education and lifestyle intervention, utilizing the Transtheoretical Model for Change, on blood lipid profile and risks for CVD in African Americans were investigated (Qian et al., 2007). Eighty-nine African Americans aged 35-75 years, 15 men and 74 women) at CVD risk participated in the program. Lipid profiles (triacylglycerols, low density lipoprotein cholesterol [LDL-C], high density lipoprotein cholesterol [HDL-C], and total cholesterol) and plasma high-sensitivity C-reactive protein (hs-CRP) concentrations were monitored before and after nutrition education intervention. Results showed a 3% reduction in body weight ($P < 0.01$);BMI ($kg/m^{2)}$ was reduced by 3.2% ($P < 0.01$). Hip circumference was reduced from baseline by 1.22% ($P < 0.05$). Other reductions includedan 8.8% reduction ($P < 0.05$) in triacylglycerols; a 5.1% increase was measured ($P < 0.01$) in HDL-C. The decreased hs-CRP level, a 68.5% reduction ($P < 0.05$) indicated that nutrition education did reduce the inflammatory processes within the human body, which might have a beneficial effect on disease reduction.

A meta analysis (386 articles) aimed at determing the effectiveness of health programs on healthy outcomes in faith-based organizations were evaluated at approximately 200 eligible institutions. Significant reduction ($P < 0.05$) in cholesteol and blood pressure levels and weight were measured. Therefore, such programs focused on primary prevention, general health maintenance and cardiovasuclar health (DeHaven et al., 2004) can bring about improvements in overall health and reduce the risk for CVD.

Lack of exercise also negatively impacts hypertension. Martin et al (2007) evaluated the characteristics of insufficiently active hypertensive African-American women using a social cognitive theory and the Transtheoretical Model to identity positive resources and areas of need to improve activity levels. According to the Transtheoreticl Model stages of change, 88.52 % of the sample was in the contemplation state. Women reported moderate levels of confidence to overcome barriers, a moderate level of confidence to use self-motivation, and reported that barriers rarely interfered with their ability to be physiclly active. The researchers concluded that physical activity interventions should focus on developing social support networks and teaching a variety of behavior strategies important to the adoption of an active lifestyle. The same group of researchers in an effort to further determine the relationship of health behavior theories with self-efficacy among insufficiently active hypertensive African-American women identified correlates associated with self-efficacy which included: overcoming barriers to physical activity; making time for activity and sticking with physical activity. The results suggested that self-efficacy is behavior specific and each measure likely provides specific information (Martin et al., 2008).

4.4 Recommended integrative CVD prevention strategies

The obvious cardiovascular disparities existing between African Americans and Whites as well as African Americans living in rural versus urban areas, clearly suggest that an integrative strategy to reduce these racial and geographical disparities is needed. An integrated strategy (addressing the entire needs of the community) for African Americans living in rural communities must take into consideration programs that will address specific demographical, geographical, cultural, educational and socioeconomic characteristics in order to be effective in reducing high blood pressure, weight and other risk factors for CVD. Sustaining change must be a part of any program aimed at improving the lives of African Americans living in the rural South. Barriers to cardiovascular health faced by African Americans living in the rural South include: lack of food availability and accessibility, inadequate intake of fruits and vegetables and other foods, lack of knowledge about the good quality of foods already in their diet, physical inactivity and lack of the resources needed to access what is needed. These barriers must and should be overcome if we are ever to close the gap relative to health disparities. Prevention strategies should be tailored to be culturally and regionally specific to meet the needs of each community.

5. Conclusion

CVD is of public health concern particularly among African Americans living in rural communities in the southern region of the United States. African American women living in rural communities are at a particularly elevated risk for CVD. Risk for CVD increased by limited education, income and neighborhood resources, which act in concert to further widen the gap in health disparities. CVD prevention within these communities requires an integrative, culturally sensitive strategy that identifies, evaluates and optimizes demographic, environmental, food, and social factors that contribute to health promotion and disease prevention. Nutrition-based strategies should consider the context of the food environment and promote the increased consumption of more healthful foods already present in the diet. Behavior based strategies should consider the potential lack of sidewalks and other facilities that may hinder physical activity. Finally, social models and compliance tools should consider the unique family and community environments, which may be passed down from one generation to the next and how this influences sustained behavior and ultimately cardiovascular health. Preventing CVD and ultimately mitigating CVD disparities in rural communities necessitates an integrative approach encompassing an open and honest dialogue between community residents and leaders, health care professionals and federal health agencies.

6. Acknowledgment

The authors would like to acknowledge the Tuskegee University Department of Food and Nutritional Sciences, College of Agriculture, Environment and Nutrition Sciences and the Tuskegee University USDA/NIFA George Washington Carver Experiment Station.

7. References

AHRQ (2005). Healthcare Disparities in Rural Areas. Selected Findings from the 2004 National Healthcare Disparities Report. AHRQ Pub No. 05-P022.

Aiello, A. E. & Kaplan, G. A. (2009). Socioeconomic Position and Inflammatory and Immune Biomarkers of Cardiovascular Disease: Applications to the Panel Study of Income Dynamics. *Biodemography and Social Biology*, Vol.22, No.2, pp. 178-205, ISSN 1948-5573.

Albert, M. A., Glynn, R. J., Buring, J., & Ridker, P. M. (2006). Impact of Traditional and Novel Risk Factors on the Relationship Between Socioeconomic Status and Incident Cardiovascular Events. *Circulation*, Vol. 114, No. 24, pp. 2619-2626, ISSN 1524-4539.

Appel, S. J., Harrell, J. S., & Deng, S. (2002). Racial and Socioeconomic Differences in Risk Factors for Cardiovascular Disease Among Southern Rural Women. *Nursing Research*, Vol. 51, No. 3, pp. 140-147, ISSN 1538-9847.

Barnett, E. & Halverson, J. A. (2000). Disparities in Premature Coronary Heart Disease Mortality by Region and Urbanicity Among Black and White Adults ages 35-64, 1985-1995. *Public Health Reports*, Vol. 115, No. 1, pp. 52-64. ISSN 0033-3549.

Barnett, E., Halverson, J. A., Elmes, G. A. & Braham, V. E. (2000). Metropolitan and Non-Metropolitan Trends in Coronary Heart Disease Mortality within Appalachia, 1980-1997. *Annals of Epidemiology*, Vol. 10, No. 4, pp. 370-379, ISSN 1407-2797.

Blumenthal, S. J. & Kagen, J. (2002). The Effects of Socioeconomic Status in Health in Rural and Urban America. Vol. 287, No. 1, pp. 109, ISSN 1538-3598.

Bovell-Benjamin, A. C., Hathorn, C. S., Ibrahim, S., Gichuhi, Bromfield, E. M. (2009). Healthy Food Choices and Physical Activity Opportunities in two Contrasting Alabama Cities. *Health and Place*. 15: 429-438.

Brown, D. L. & Warner, M. E. (1991). Persistent Low-Income Nonmetropolitan Areas in the United States: Some Conceptual Challenges for Policy Development. *Policy Studies Journal*, Vol. 19, No. 2, pp. 22-41, ISSN 1541-0072.

Brown, M. J. (2006). Hypertension and Ethnic Group. *British Medical Journal*, Vol. 332, No. 7545, pp. 833-839, ISSN 1020-9077.

Brownson, R. C., Smith, C. A., Patt, M., Mack,N. E.,Thompson,J. J., Dean, C. G., Dabney, S. & Wilkerson, J. C. (1996). Preventing Cardiovascular Disease through Community-Based Risk Reduction: The Bootheel Heart Health project. *Amercian Journal of Public Health*. Vol. 86, No. 2, pp. 206-212, ISSN 1541-0048.

Brownson, R. C., Baker, E. A., Boyd, R. L., Caito, N. M., Duggan, K. Et al. (2004). A community-based Approach to Promoting Walking in Rural Areas. *American Journal of Preventive Medicine* 27: 28-34. DOI.10. 1016

Bureau of Economic Analysis. U. S. Department of Commerce. State Personal Income 2010. http://www.bea.gov/newsreleases/regional/spi/2011/pdf/spi0311.pdf

Carter, V. L., Dawkins, N. L., & Howard, B. (2009) Addressing Health Disparities through Cancer Awareeness abnd Prevention. *UnPublished Data*

CDC (2010). Heart Disease Death Rates 2000-2006. Accessed October, 12, 2011

CDC, 2011. Healthy People, 2000, 2010 , 2020. Accessed October 1, 2011

CDC Healthy People 2000, Final Review,Department of Health & Human Services, Publication No. 01-0256,Hyattsville, MD.

Centers for Disease Control and prevention, US. Obesity Trend, Centers for Disease Control and Prevention, Atlanta, GA. Accessed October, 12, 2011.

Collins, R. (2002) African American Women and Men at High and Low Risk for Hypertension a Signal Detection and Analyses of NHANES III, 1988-1994. *Preventive Medicine*, Vol.35, No. 4, pp. 303-312, ISSN 0091-7435.

Cooper, R., Cutler, J., Desvigne-Nickens, P., et al. (2000). Trends and Disparities in Coronary Heart Disease, Stroke, and Other Cardiovascular Diseases in the United States: Findings of the National Conference on Cardiovascular Disease Prevention. *Circulation*, Vol. 102, No. 25, pp. 3137-3147, ISSN 1524-4539.

Cort, N. A. & Stewart-Fahs, P. (2001). Heart Disease: The Hidden Killer of Rural Black Women. *Journal of Multicultural Nursing & Health*, Summer 2001, Available from http://findarticles.com/p/articles/mi_qa3919/is_200107/ai_n8962518/pg_4/?tag=mantle_skin;content.

Danaei, G., Rimm, E. B., Oza, S., Kulkarni, S. C., Murray, C. J. L., et al. (2010). The Promise of Prevention: The Effects of Four Preventable Risk Factors on National Life

Expectancy and Life Expectancy Disparities by Race and County in the United States. *Public Library of Science Medicine*, Vol. 7, No. 3, pp. e1000248. doi:10.1371/journal.pmed.100024.

Dawkins, N. D. , Ward, J. A., Shikany, J. & Pace, R. D. (April/May 2009). Boost for Purslane. *The World of Food Ingredients*. pp. 58-60.

DeHaven, M. J., Hunter, I. B., Wilder, L., Walton, J. W., & Berry, J. (2004). Health Program in Faith Based Organization : Are They Effective? *American Journal of Public Health*, Vol. 94, No. 6, pp. 1030-1036, ISSN 1541-0048.

Division for Heart Disease and Stroke Prevention: Data Trends & Maps Web site. U. S.

Department of Health and Human Services, Centers for Disease Control and Prevention (CDC), National Center for Chronic Disease Prevention and Health Promotion, Atlanta, GA, 2010. Available at http://www.cdc.gov/dhdsp/.

Centers for Disease Control and Prevention (CDC). *Behavioral Risk Factor Surveillance System Survey Data.* Atlanta, Georgia: U. S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2010. http://apps.nccd.cdc.gov/brfss/.

Dunn, R. A., Sharket, J. R., Lotade-Manje, J., Bouhlal, Y. & Nayga, R. M. (2011). Socio-economic Status, Racial Composition and The Affordability of Fresh Fruits and Vegetables in Neighborhoods of a Large Rural Region in Texas. *Nutrition Journal*, Vol. 10, No. 6, ISSN 1475-2891.doi:10.1186/1475-2891-10-6.

Eberhardt, M., Ingram, D., Makuc, D. M., et al. (2001). Urban and Rural Health Chartbook. Health, United States, 2001. Hyattsville, Maryland : National Center for Health Statistics. Available from http://cdc.gov/nchs/data/hus/hus01cht_ac.pdf.

Eberhardt, M. S. & Pamuk, E. R. (2004). The Importance of Place of Residence: Examining Health in Rural and Nonrural Areas. *American Journal of Public Health*, Vol. 94, No. 10, pp. 1682-1686, ISSN 1541-0048.

Elyer. A. A., Koffman, D. M., Young, D. R., Wilcox, S. Wilbur, J. Eta al. (2003). Auantitative study of correlates of physical activity in Women for Diverse Racial/Ethnic Groups: The Women's Cardiovascular Health Network Project Summary and Conclusion. *American Journal of Preventive Medicine,* Vol. 25, No. 3, pp.93-103, ISSN0749

Fakiri, F. E., Bruijnzeels, M. A. & Hoes, A. W. (2006). Prevention of Cardiovascular Diseases: Focus on Modifiable Cardiovascular Risk. *Heart*, Vol. 92, No.. 6, pp. 741-745, ISSN 1468-201X.

Federal statistics. (2004) Accessed September, 2009.
www. Fedstats.gov/gf/states/01/01/087.html

Fincher, C.., Williams, J. E., MacLean, V., Allison, J. J.., Kiefe C. I. et al. (2004) Racial Disparities in Coronary Heart DFiseae: A Sociological View of the Medical Literature on Physician Bias. Ethnicity & Disease, Vol. 14, Summer,pp. 360-371,, ISSN 1049-510X

Ford, C. D., Kim, M. J., & Dancy, B. L. (2009). Perceptions of Hypertension and Contributing Personal and Environmental Factors Among Rural Southern African American Women. *Ethnicity and Disease*, Vol. 19, No. 4, pp. 407-413, ISSN 1049-510X.

Fullerton, H. J. Risk of stroke in children: Ethinic and Gender Disparities. (2003). *Neurology* Vol. 61, No. 2, pp. 189-194, ISSN 1526-632X .

Galbardes, B., Smith, G. D., & Lynch, J. W. (2006). Systematic Review of the Influence of Childhood Socioeconomic Circumstances on Risk for Cardiovascular Disease in Adulthood. *Annals of Epidemiology*, Vol. 16, No. 2, pp. 91-104, ISSN 1047-2797.

Gillum, F. R., (1999). Stroke Mortality in Blacks: Distrubing Trends. *Stroke*, Vol. 30, No. 8, pp. 1711-1715, ISSN 1524-4628.

Glover, S., Piper, C. N., Hassan, R., et al., (2011). Dietary, Physical Activity, and Lifestyle Behaviors of Rural African American South Carolina Children. *Journal of the National Medical Association*, Vol. 103, No. 4,pp. 300-304, ISSN 0027-9684.

Griswold, D. M. (1927). Heart Disease As A Public Health Problem. *American Journal of Public Health*, Vol. 17, No. 2, pp.101-107, ISSN 1541-0048.

Hackman, D. G. & Anand, S. S. (2003). Emerging Risk Factors for Atherosclerotic Vascular Disease : A Critical Review of the Evidence. *Journal of the American Medical Association*, Vol. 290, No. 7, pp. 932-940, ISSN 1538-3598.

Hajjar, I. & Kotchen T. J. (2003). Regional Variations of Blood Pressure in the United StatesAre Associated with Regional Variations in Dietary Intakes: The NHANES-III Data. *Journal of Nutrition*, Vol. 133, No. 1, pp. 211–214, ISSN 1541-6100.

Heidenreich, P. A., Trogdon, J. G., Khavjou, O. A. et al. (2011). Forecasting The Future of Cardiovascular Disease in the United States : A Policy Statement From The American Heart Association. *Circulation*, Vol. 123, No. 8, pp. 933-944, ISSN 1524-4539.

Hermstad, A. K., Swan, D. W., Kegler, M. C., Barnette, J. K. & Glanz, K. (2010). Individual and Environmental Correlates of Dietary Fat Intake in Rural Communities: A Structural Equation Model Analysis. *Social Science and Medicine*, Vo. 71, No. 1, pp. 93-101, ISSN 0277-9536.

Hertz, R. P., Unger, A. N., Cornell, J. A. & Saunders, E. (2005). Racial Disparities in Hypertension Prevalence, Awareness and Management. *Archives of Internal Medicine*, Vol. 165, No. 18, pp.2098-2104, ISSN 0003-9926.

Hosler, A. S., Rajulu, D. T., Fredrick, B. L., & Ronsani, A. E. (2008). Assessing Retail Fruit and Vegetable Availability in Urban and Rural Underserved Communities. *Preventing Chronic Disease: Public Health Research, Practice and Policy*, Vol. 5, No. 4, pp. 1-9, ISSN 1545-1151.

Huang, Z., Wang, B., Pace, R. D., Yoon, S. (2009). Identification of Anthocyanins in Muscadine Grapes with HPLC-ESI-MS. *Food Science and Technology,*Vol. 42, No. 4,pp. 819-824, ISSN 0022-1155.

Huang, Z., Wang, B. &Pace, R. D.(2008). Potentials of Phytochemicals in Disease Prevention and Health Promotion, In:*Recent Advances in Agriculture*, Stevens, C. & Khan, V. A., pp. 63-97, Research Signpost, ISBN 978-81-308-0222-0, Kerala, India.

Huang Z., Wang, B., Eaves, D. H., Shikany, J. M., Pace, R. D. (2007). Phenolic Compound Profiles of Selected Vegetables Frequently Consumed by African Americans in the Southeast United States. The ACCE Project. *Journal of Food Chemistry,* Vol. 103, No. 4, pp. 1395-1402, ISSN 0308-8146.

Huang, Z., Wang, B., Eaves, D. H., Shikany, J. M. &Pace, R. D.(2007). Total Phenolics and AntioxidantCapacity of Indigenous Vegetables in The Southeast United States: The ACCE Project. *International Journal of Food Science and Nutrition*, Vol. 60, No.2, pp. 100-108, ISSN 0963-7486.

Jackson, E. J., Doescher, M. P., Jerant, A. F & Hart, L. C. (2005). A National Study of Obesity Prevalence and Trends by Type of Rural County. *The Journal of Rural Health,*Vol. 21, No. 2, pp. 140-148, ISSN 1748-0361

Johnson, M. and Pace, R. D. (2010). Sweet Potato Leaves: Properties and Synergistic Interactions That Promote Health and Prevent Disease. *Nutrition Reviews*, Vol. 68, No. 10, pp. 604-615, ISSN 1753-4887.

Johnson, K. (2006). Demographic Trends in Rural and Small Town America. *Reports on Rural America*, Vol. 1, No. 1, Carsey Institute, University of New Hampshire, Durham, New Hampshire.

Kaufman, J. Childhood Obesity Statistics and Facts.2007. National Association of Children's Hospital and Related Institutions and National Association Children's Hospitals, public policy affiliate. Accessed January 2010

Kennedy, B. S. (2002). Trends in Hospitalized Stroke for Blacks and Whites in the United States,1980-1999. *Neuroepidemiology*, Vol. 21, No. 3, pp. 131-141, ISSN 1423-0208.

Khan, J. A. , Casper, M., Asimos, A. W., Clarkson, L., Enright, D., Fehrs, L. J., George, M., Heidari, K., Huston, S. L., Mettam, L. H., Williams, I., Schieb, L., & Greer, S. (2011). Geographic and Sociodemographic Disparities in Drive Times to Joint Commission-Certified Primary Stroke Centers in North Carolina, South Carolina and Georgia. *Preventing Chronic Disease : Public Health Research, Practice and Policy*, Vol. 8, No. 4, A79, pp. 1-8, ISSN 1545-1151.

Khot, U. N., Khot, M. B., Bajzer, C. T., Sapp, S. K., Ohman, E. M., Brener, S. J., Ellis, S. G., Lincoff, A. M. & Topol, E. J. (2003). Prevalence of Conventional Risk Factors in Patients With Coronary Heart Disease. *Journal of the American Medical Association*, Vol. 290, No. 7, pp. 898-904, ISSN 1538-3598.

Larson, S. L. & Fleishman, J. A. (2003). Rural-Urban Differences in Usual Source of Care and Ambulatory Service Use: Analyses of National Data Using Urban Influence Codes. *Medical Care*, Vol. 41, No. 7 Suppl, pp. III 65-III74, ISSN 1537-1948.

LaVonna, B. L, Sloane, D. C., Nascimento, L. M., Diamant, A. A. L., Guinyard, J. J., Yancey, K., & Flynn, G. (2005). African Americans' Access to Healthy Food Options in South Los Angeles Restaurants. *American Journal of Public Health*, Vol. 95, No. 4, pp 668-673. ISSN 1541-0048.

Lewis, L.B., Sloane, D. C., Nascimento, L. M., Diamant A. L., Guinyard, J. J.,Yancey, A. K., Flynn, G. (2005). African Americans' Access to Healthy Food Options in South Los Angeles Restaurants. *Am. J. Public Health* 95: 668-673. DOI.10.2105.

Lloyd-Jones, D., Adams, R., Carnethon, M., et al., (2009). Heart Disease and Stroke Statistics_2009 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, Vol. 119, No. 3, pp. e21-e181, ISSN 1524-4539.

Lloyd-Jones, D. M., Hong, Y., Labarthe, D., et al., (2010). Defining and setting national goals for cardiovascular health promotion and disease reduction : the American Heart Association's strategic impact goal through 2020 and beyond. *Circulation*, Vol. 121, No. 4, pp.586-613, ISSN 1524-4539.

Loughlin, J. L., Paradis, G., Donald, K. G. & Renaud, L., (1999). *Amercial Journal of Public Health, Vol.89*: No 12, pp. 1819-1826, ISSN 1541-0048.

Lyons, M. A., Crow, C., Dunn, L., Edwards, B., Graves, A. Shelton, M. & Dunkin, J. (2001). Partnership for Healthier Rural Communities. *Online Journal of Rural Nursing and Health Care*, Vol. 2, No. 2, pp.38 -43, ISSN 1539-3399.

Mainous, A. G, King, D. E., Garr, D. R. & Pearson, W. S. (2004). Race Rural Residence, and Control of Diabetes and Hypertension. *Annals of Family Medicine*, Vol. 2, No. 6,pp. 563-568, ISSN 1544-1717.

Martin, M. Y. Prayer-Patterson, H, Kratt, P, Kim, Y, Person, S. (2007). Characteristics of Insufficiently Active Hypertensive Black Women who Volunteer to be in a Physical Activity Promotion Intervention:An application of social Cognitive Theory and the Transtheroretical Model. Ethnicity and Disease. 17 (4) 604-610, ISSN 1049-510X.

Martin, M. Y. Prayer-Patterson, H, Kratt, P, Kim, Y, Person, S. (2007). Characteristics of Insufficiently Active Hypertensive Black Women who Volunteer to be in a Physical Activity Promotion Intervention:An application of social Cognitive Theory and the Transtheroretical Model. Ethnicity and Disease. 17 (4) 604-610, ISSN 1049-510X.

McGruder, H. F., Malarcher, A. M., Atoine, T. L., Greenlund, K. J. & Croft, J. B. (2004). Racial and Ethnic Disparities in Cardiovascular Risk Factors Among Stroke Survivors. *Stroke*, Vol. 35, No. 7 , pp. 1557-1561, ISSN 1524-4628.

Mueller, K. J., Ortega, S. T., Parker, K. & Patil, K. (1999). Health Status and Access to Care Among Rural Minorities. *Journal of Health Care for the Poor and Underserved*, Vol. 10, No. 2, pp. 230-249, ISSN 1049-2089.

Muennig, P., Sohler, N. & Mahato, B. (2007). Socioeconomic Status As An Independent Predictor of Physiological Biomarkers of Cardiovascular Disease: Evidence from NHANES. *Preventive Medicine*, Vol. 45, No. 1, pp. 35-40, ISSN 0091-7435.

Nutrition Health and Physical Activity Fair (NHPAF)-The 2008, 2009, 2010 and 2011. Department of Food and Nutritional Sciences, Tuskegee University, Tuskegee, AL.

Ogden, C. L., Carroll, M. D., Curtain, L. R., McDowell, M. A., Tabak, C. J., & Flegal, K. M. (2006). Prevalence of Overweight and Obesity in the United States, 1999-2004. *Journal of The American Medical Association*, 295: 1549-1555. Retrieved October 14,2010 from: http://www.cdc.gov/healthyyouth/overweight/index.htm

Ong, K. Cheung, Y. B., Man, C. P.,Lau, K. & Lam, S. K. (2007). Prevalence, Awareness, Treatment, and Control of Hypertension among United States Adults 199-2004. *Hypertension*, Vol.49, pp. 69-75, ISSN 0194-911X.

Pace, R. D., Dawkins, N., Wang, B., Person, S. & Shikany, J. M. (2008). Rural African Americans' Dietary Knowledge, Perceptions, and Behavior in Relation to Cardiovascular Disease. *Ethnicity & Disease*, Vol. 18, No. 1, pp. 6-12, ISSN 1049-510X.

Pearson, T. A. & Lewis, C. (1998). Rural Epidemiology: Insights from a Rural Population Laboratory. *AmericanJournal of Epidemiology*, Vol. 148, No. 10, pp. 949-957, ISSN 1476-6256.

Qian, J., Wang, B., Dawkins, N. L., Gray, A. &Pace, R. D. (2007) Reduction of Risk Factors for Cardiovascular Diseases in African Americans with a 12-week Nutrition Education Program. *Nutrition Research*, Vol. 27, No. 5, pp. 252-257, ISSN 0271-5317.

Ranjit, N., Diez-Roux, A. V., Shea, S., Cushman, M., Ni, H., Seeman, T. (2007). Socioeconomic Position, Race/Ethnicity, and Inflammation in the Multi-Ethnic Study of Atherosclerosis. *Circulation*, Vol. 116, No. 21, pp. 2383-2390, ISSN 1524-4539.

Rankins, J., Sampson, W. Brown, & Jenkins-Sally, T. (2005). Dietary Approaches to Stop Hypertension (DASH) Intervention Reduces Blood Presure among Hypertensive African American Patients in a Neighborhood Health Care Center. *Journal of Nutrition Education and Behavior*, Vol. 37, No. 5, pp. 259-264, ISSN 1499-4046.

Resnicow, K. K.,Campbell, C., Carr, F.,Mccarty, T., Wang, S.et al. (2004). A dietary Intervention Conducted Through African American Churches. American Journal of Preventive Medicine, Vol 27, No.2, pp. 97-105, ISSN 0749-3797.

Reschovsky, J. D. & Staiti, A. B. (2005). Access and Quality: Does Rural America Lag Behind? *Health Affairs*, Vol. 24, No. 4, pp. 1128-1139, ISSN 1544-5208.

Reynnells, L. & John, P. L. (2008). What is Rural? Rural Information Center, National Agricultural Library, U. S. Department of Africulture, Beltsville, MD.

Roger, V. L., Go, A. S., Lloyd-Jones, D. M., et al. (2011). Executive Summary: Heart Disease and Stroke Statistics—2011 Update: A Report From the American Heart Association. *Circulation*, Vol. 123, No. 4, pp. 459-463, ISSN 1524-4539.

Rosvall, M., Engström, G., Janzon, L., Berglund, G. & Hedblad, B. (2007). The role of low grade inflammation as measured by C-reactive protein levels in the explanation of socioeconomic differences in carotid atherosclerosis. *European Journal of Public Health*, Vol. 17, No. 4, pp. 340-347, ISSN 1464-360X.

Sanderson, B. K., Foushee, H. R., Bittner, V., et al. (2003). Personal, Social, and Physical Environmental Correlates of Physical Activity in Rural African-American Women in Alabama. *American Journal of Preventive Medicine*, Vol. 25, No. 3Si, pp. 30-37, ISSN 0749-3797.

Savoca, M. R., Arcury, T. A., Leng, X. (2009). The Diet Quality of Rural Older Adults in the South as Measured by Healthy Eating Index-2005 Varies by Ethnicity. *Journal of the American Dietetic Association*, Vol. 190, No. 12, pp. 2063-2067, ISSN 0002-8223.

Sharkey, J. R., Johnson, C. M. & Dean, W. R. (2010). Food Access and Perceptions of the Community and Household Food Environments as Correlates of Fruit and Vegetable Intake Among Rural Seniors. *BMC Geriatrics*, 10: 32, ISSN 1471-2318. http://www.biomedcentral.com/1471-2318/10/32.

Stampfer, M. L., Hu, F. B., Manson, J. E., Rimm, E. B., & Willett, W. C. (2000). Primary Prevention of Coronary heart Disease in Women Throught Diet & Lifestyle. *The Journal of Medicine*, Vol. 343, No. 1, pp.17-22, ISSN 0028-4793.

Sumner, A. E. & Cowie, C. C. (2008). Ethnic Difference in the Ability of Triglyceride Levels to Identify Insulin Resistance. *Atherosclerosis*, Vol. 196, No. 2, pp.696-703, ISSN 0026-0495.

Sumner, A. E., Finley, K. B., Genovese, D. J., Criqui, M. H. & Boston, R. C. (2005a). Fasting Triglyceride and the Triglyceride–HDL Cholesterol Ratio Are Not Markers of Insulin Resistance in African Americans. *Archives of Internal Medicine*, Vol. 165, No. 12, pp. 1395-1400, ISSN 0003-9926.

Sumner, A. E., Vega, G. L., Genovese, D. J., Finley, K. B., Bergman, R. N. & Boston, R. C. (2005b). Normal Triglyceride Levels Despite Insulin Resistance in African Americans: Role of Lipoprotein Lipase. *Metabolism*, Vol. 54, No. 7, pp.902-909, ISSN 0026-0495.

Tabassum, F. M., Kumari, M., Rumley, A., Lowe, G., Power, C. & Strachan, D. P. (2008). Effects of Socioeconomic Position on Inflammatory and Hemostatic Markers: A Life-Course Analysis in the 1958 British Birth Cohort. *American Journal of Epidemiology*, Vol. 167, No. 11, pp. 1332-1341, ISSN 0002-9262.

Taylor, H. A., Hughes, G. D., & Garrison, R. J. (2002). Cardiovascular Disease Among Women Residing in Rural America: Epidemiology, Explantations, and Challenges. *American Journal of Public Health*, Vol. 92, No. 4, pp. 548-551. ISSN 1541-0048.

U. S. Census Bureau: State and County QuickFacts. Data derived from Population Estimates, American Community Survey, Census of Population and Housing, Small Area Income and Poverty Estimates, State and County Housing Unit Estimates, County Business Patterns, Nonemployer Statistics, Economic Census, Survey of Business Owners, Building Permits, Consolidated Federal Funds Report. http://quickfacts.census.gov/qfd/states/00000.html. Accessed October 20, 2011

U. S. Census Bureau. State and Metropolitian Area Data Book: 2010. Retrieved October 19, 2011. http://www.census.gov/prod/www/abs/smadb.html.

Wang Y, and Beydoun MA. The Obesity Epidemic in the United States—Gender, Age, Socioeconomic, Racial/Ethnic, and Geographic Characteristics: A Systematic Review and Meta-Regression Analysis. *Epidemiology Review* 29 (1) 6-28.

Wilcox,S. , Laken, M., Anderson, T.,Bopp, M., Bryant, D. et al. The health e-AME faith-based Physical Activity Initiative: description and Baseline Findings. *Health promotion Practice*, Vol 8, pp. 69

Wilcox, S., Bopp, M., Oberrecht, L., Kammermann, S. K. & McElmurray, C. T. (2003). Psychosocial and Perceived Environmental Correlates of Physical Activity in Rural and Older African American and White Women. *Journal of Gerontology: Physiological Sciences*, Vol. 58B, No. 6, pp. P329-P337, ISSN 1758-5368.

Wilson, T. W., Kaplan, G. A., Kauhanen, J., Cohen, R. D., Wu, M., Salonen, R. & Salonen, J. T. (1993). Association between Plasma Fibrinogen Concentration and Five Socioeconomic Indices in the Kuopio Ischemic Heart Disease Risk Factor Study. *American Journal of Epidemiology*, Vol. 137, No. 3, pp. 292-300, ISSN 1332-1341.

Wimberley, R. C. & Morris, L. V. (1997). The southern black belt: A national perspective. Lexington, KY TVA Rural Studies (The University of Kentucky). USA

Winham, D. M., & Jones, K. M. (2011). Knowledge of Young African American Adults About Heart Disease: A Cross-Sectional Survey. *BMC Public Health*, Vol. 11, No. 248, pp. 248-58, ISSN 1471-2458.

Winkleby, M. A., Jatulis, D. E., Frank, E. & Fortmann, S. P. (1992). Socioeconomic Status and Health : How Education, Income and Occupation Contribute to Risk Factors for Cardiovascular Disease. *American Journal of Public Health*, Vol. 82, No. 6, pp. 816-820, ISSN 1541-0048.

World Health Organization (WHO) 2011. Heart disease. Accessed Oactober, 1 2011

Yanek, L. R., Becker, D. M. Moy, T. F., Gittelsohn, J. & Koffman, D. M. (2001). Project Joy: Faith Based Cardiovascular Health Promotion for African American Women. *Public Health Reports*, Vol. 116, No. Suppl 1, pp. 68-81, ISSN 0033-3549.

Yarnell, J., McCrum, E., Arveiler, D., et al., (2005). Education, Socioeconomic and Lifestyle Factors, and Risk of Coronary Heart Disease: The PRIME Study. *International Journal of Epidemiology*, Vol. 34, No. 2, pp. 268-275, ISSN 1464-3685.

Zuniga, M., Anderson, D. A. & Alexander, K. (2003). Heart Disease and Stroke in Rural America: A Literature Review. Rural Healthy People 2010 : A Companion Document to Healthy People 2010. Volume 2. College Station, TX : The Texas A&M University System Health Science Center, Southwest Rural Health Research Center.

4

Gender Differences in Food Choice and Dietary Intake in Modern Western Societies

Claudia Arganini, Anna Saba,
Raffaella Comitato, Fabio Virgili and Aida Turrini
*INRAN - Istituto Nazionale di Ricerca per gli Alimenti e la Nutrizione
(National Research Institute for Food and Nutrition), Rome,
Italy*

> *"Sexual difference is probably the issue in our time
> which could be our 'salvation' if we thought it through."
> Luce Irigaray*

1. Introduction

A significant "male oriented" bias in science is a matter of fact (Marino et al., 2011), even though the number of women majoring in science has increased dramatically (Yokoo, 1996). Considering people graduated in mathematics, science and technology per 1,000 of population aged 20-29, since 1993 up to 2009 (EUROSTAT, 2011), proportion of women is 4.4% vs. 3.4% of men. Interestingly, 4 out of 5 authors of this paper are women. It is justifiable to ask the reason for this gender bias. We can advocate two "reasonable" reasons: the first one has mainly sociologic concern in that different aspects of the human society still present an odd distribution. Research is not an exception and even though the proportion of women within professionals involved in public and private research has reached and possibly overtaken the other gender, the target of scientific investigations is still to be actually considered somehow unbalanced in favour of one gender. The second one has an exquisitely pragmatic origin: in the majority of cases, and unless your research is not to be focused on events strictly connected to females (pregnancy, lactation, few organ specific disease), males are a simpler and cheaper experimental model than females. No needs to carefully evaluate risk factors bound to pregnancy and lactation, to consider hormonal cycle, no sharp changes of tissue functionality associated with ageing.

It is a matter of fact that in science, and nutritional science is not an exception, there is a widely accepted overlapping between the terms "human" and "male", while the term female (or woman in the case of the specie H. Sapiens) strictly refers to "not males". At the same time, the majority of us would agree in defending the evidence that both the biochemistry and the physiology significantly differ between genders, even independently on the most evident female physiological characteristics of presenting a specific and cyclic exposure to hormones flux. These differences are consistent with a gender-specific genetic set up, and result in a specific capacity to relate to and cope with the environmental challenge.

Recent literature (Marino et al., 2011) has highlighted that nutrition could differently influence the health of male and female individuals. It is widely accepted that nutrition is not only "just a fuel" but is the most significant part of the environment that we actually introduce into our body and eating patterns are a relevant component of the cultural reference models (Harris, 1985).

Driven by the above considerations this chapter will present and discuss available data emerging from an extensive literature review addressing differences and similarities between genders in food choice and food consumption patterns in modern western societies. An original elaboration of data on food consumption profiles according to gender will also be presented and critically evaluated under the perspective of nutrients intake and fulfilment of nutritional requirements at population level. In the preparation of this overview, we have undertaken a bibliographical search limited to social and scientific literature published in English. The search, informed by a strongly limited selection of words, included databases of peer-reviewed literature (SCOPUS) from 1995 to 2011. A 'snowball procedure' was employed whereby the references cited in each article were browsed for further relevant research. An original elaboration of data on food consumption profiles according to gender will also be presented and critically evaluated under the perspective of nutrients intake and fulfilment of nutritional requirements.

2. Key determinants in food choice: A gender perspective

Food choices is an area in which research has revealed consistent behavioural gender differences. Food choice is dependent on a wide spectrum of factors, which affect human behaviour in different ways, resulting alternatively in the choice of some specific products and in the rejection of others. The study of food choice is mostly dealing with one question: "why do people eat the foods they eat?" Food plays an important part in all our lives in a variety of ways. The choices people make among foods determine which nutrients enter the body. However, in modern societies, food is more than mere sustenance. What people choose to eat is not solely based on their biological needs, their choice also addresses many psychological and/or emotional issues (Conner & Armitage, 2002). After all, a person does not necessarily have to be hungry to eat, does not always choose his/her most preferred food, and some of the influences in food choice might be unconscious. Generally speaking, food choice is a complex human behaviour and consequently is influenced by many interrelating factors ranging from biological mechanism and genetic profiles to social and cultural factors. Many studies have explored selected aspects of food choices from an ample variety of disciplines and perspectives (Axelson & Brinberg, 1989; Booth, 1994; Glanz et al., 1992; Mennell et al., 1992; Murcott, 1983; Shepherd, 1990, 2005; Thompson, 1988). Recent notions generally split the factors influencing food choice into those related to the food, to the person making the choice and to the external economic and social context in which the choice is made (Booth and Shepherd, 1988; Randall and Sanjur, 1981). There are chemical components and physical properties of the food which are likely to have an impact on choice, via sensory perception. However, perceiving a sensory attribute in a food does not necessarily means that a person will choose to consume that food. It is the person's liking for that specific attribute in that food which influences choice. Psychological differences between people, such as personality, may also influence food choice. In addition to factors associated with the person and the food, there are also other many factors in the context within which the choice is made that can be important in food choice. These include marketing and economic variables as well as social, cultural, religious or demographic variable. Food choices are made by individuals from alternatives available in a certain use situation. They are made repetitively, every day in various use situations: what to buy and take home for the family, what to eat at the canteen and which dishes to choose from a menu when eating out at a restaurant. Food choice may also be characterized by the context, a situation determined by the time, place, and company. In Western societies the abundance and variety of foods to choose is extensive. Anthropological and sociological work has emphasized the meaning of food and eating in self and cultural definition (Berbesque, 2009; Counihan, 1999; Murcott, 1983; Vartaniana et al., 2007). This literature indicates that, as mentioned above, the importance of food and eating extends well beyond a the need of covering "physiological needs", playing a role in identity expression, communication, social interactions, as well as in delineating status and gender roles. Eating behaviour is therefore likely to be vulnerable to various social influences, including the desire to respond in a socially-desirable manner (Herman et al., 2003). Studies by Lindeman and colleagues (Lindeman & Sirelius, 2001; Lindeman & Stark, 1999, 2000) suggest that food choice is a means by which one expresses one's own philosophy of life. In addition, the current emphasis on dieting and slimness in Western cultures promotes norms describing "what and when" one should eat, as well as what one should look like. Taken together, these considerations suggest that what one eats has important implications for social judgments. In addition, social changes such as the increased participation of women in the workforce lead to reduced time available for food selection and meal preparation, which further complicates food choice. Contemporary consumers have fears and conflicts involving food and health (Mennell et al., 1992; Rozin et al., 1999), and social norms about foods and meal composition, that guided previous generations, appear to be eroding, leaving people with a lack of structure related to food and eating behaviour (Fischler, 1980).

A body of literature has consistently found that many variables may influence eating behaviour, but their interrelations make their effect difficult to distinguish. In addition, the analysis of the effects of single or multiple factors is further complicated by the fact that eating behaviour is not a constant phenomenon, but will change with differing circumstances and experiences of the individual.

Studies conducted in modern western societies report consistent associations between gender and specific foods, where meat (especially red meat), alcohol, and hearty portion sizes are associated with masculinity, while vegetables, fruit, fish and sour dairy products (e.g., yogurt, cottage cheese) are associated with femininity (Jensen & Holm, 1999; Sobal, 2005). The results of a study conducted on the Hazda, a tribe of human foragers living in Tanzania, also showed a sex differences in food preferences, with males preferring meat more and females preferring berries more (Berbesque, 2009). Overall, the most relevant differences according to gender in food choices in modern western societies, emerging from our literature review, were in the relationship between eating habits and health consciousness, and between eating behaviour and weight control. Those topics will be discussed in the following sections.

2.1 Eating habits and health consciousness

In general, women have been frequently reported to engage in far more health-promoting behaviours than men and have healthier lifestyle patterns (Courtenay, 1998, 2000; Gough & Conner, 2006; Kandrack et al., 1991; Lonnquist et al., 1992; Roos et al., 2001).

Men usually talk about eating as habitual and routine, and as necessary activity to "fuel" their "fleshes". Although they are aware of "healthy eating guidelines", they often show skepticism and resistance to nutrition education messages, and frequently perceive healthy eating as monotonous and unsatisfying. Some men do express interest in food, cooking, and health, and indicate that they are reducing their consumption of red meat and increasing consumption of vegetables (Sobal, 2005). These alternative experiences with food are more commonly expressed by "high educational levels", such as engineers, than by "blue-collars workers, such as carpenters or drivers, suggesting that social class may mediate associations between "masculinity" and food (Roos et al., 2001; Sobal, 2005).

With regard to eating habits, a large number of reports indicate that in general, women are more aware about diet and health-diet relationship implications and also embrace suggested dietary changes to a greater degree than men (Barker et al., 1995; Courtenay, 2000; Friel et al., 1999; Girois et al., 2001; Thiele & Weiss, 2003). Data on a representative survey in the Norwegian population (Fagerli & Wandel, 1999) shows that women considered health aspects and chose accordingly the foods they consider to be healthy, more often than men when selecting foods for an everyday dinner. Accordingly, their reported changes more often are in line and agree with dietary guidelines. The same study also reported consistent associations of healthier food behaviours with increased age, higher education, and female gender. These findings are similar to the observations resulting from the analyses from a population data set conducted in 114 worksites in the USA, overall employing 37,291 workers who were engaged in a variety of activities (Hunt et al., 1997). Also in this study, female gender was associated with food choices closest to the recommendations to increase fiber, fruits and vegetables and to reduce fat. A single exception was in found the adherence to follow the recommendation to increase consumption of beans and lentils for which male gender were associated with greater consumption. In a Pan-EU survey of 14331 subjects, female respondents perceived that "quality/freshness", "price", "trying to eat healthy" and "family preferences" were the most important influences affecting food choice, whereas "taste" was the most frequently selected factor affecting food choice of male respondents (Lennenäs et al., 1997). In a different study, females have been reported to be more likely than males to mention more vegetables or less fat or balance as a part of a healthy diet (Margetts et al., 1997).

Another factor contributing to food choices is the persuasion by others or by specific circumstances. More women than men reported that influence of other people can prevent them from eating healthier (Lappalainen et al., 1997). It has also been shown that men give lower priority to health compared to other considerations, such as taste and convenience, in making their food choices (Steptoe et al., 2002; Wardle and Griffith, 2001) and that they feel more ambivalent about healthy dietary choices (Povey et al., 2001; Sparks et al, 2001). Earlier studies have found significant gender differences in opinions and behaviour with regard to different health aspects. For instance they reported that men choose fewer high-fiber foods, eat fewer fruits and vegetables and low-fat foods, and consume more soft drink that women

(Beardsworth et al., 2002; Beer-Borst et al., 2000; Fulkerson et al., 2004;Li et al., 2000; Liebman et al., 2001; Pollard et al., 2002; Prättälä et al., 2007; Wardle et al., 2000). The International Health and Behaviour survey (IHBS) examined a range of health behaviours in a total of 19298 university students from 23 different countries utilizing a study approach based on a self-report questionnaire (Wardle et al, 2004). In almost all of the 23 countries a higher percentage of women reported to avoid high fat-foods, to eat fiber-rich foods, to eat fruit daily. Differences in salt intakes were less consistent but nevertheless a significant female advantages in 6 countries was observed. Similarly, in a study on 1024 UK adults, women reported to consume a larger number of portion of fruit and vegetables, than men (Baker & Wardle, 2003). A research carried out in the UK, the Netherlands and in Finland documents that women are more interested than men in eating healthily and natural products (Roininen et al., 2001). The food choice motivations of a representative sample of 9339 Polish respondents depended mostly on gender and age (Wadolowska et al., 2008). The study confirmed the findings of other authors about the role of females and its correlation with health-concerned attitudes, inclination to comply with dietary recommendations and readiness to gain new nutrition knowledge. In a nationally representative sample of Irish adults (n=1256), it has been observed that young lowest social class, primary level education males, were the subgroup most likely to have negative attitudes or motivation towards healthy eating (Kearney et al.,2001). Studied conducted in Ireland reported that women were generally more prone to make conscious efforts to try to eat a healthy diet 'most of the time', while men were three times more likely to 'hardly ever' make such conscious efforts to eat a healthy diet (Kearney et al., 2001; Hearty et al. 2007). Data from a representative sample of 98733 Canadians (Canadian Community Health Survey) indicates that gender plays an important role in determining food choices. Women are more likely than men to choose or avoid foods following to concerns about health and, accordingly, choose or avoid foods due to their contents (Ree et al. 2008).

In general, women have been shown to be more thoughtful about food and health issues and they seem to have more moral and ecological misgivings about eating certain foods than men, who are more confident and demonstrate a rather uncritical and traditional adherence to eating profiles and pattern (Beardsworth et al., 2002; Teratanavat & Hooker, 2006; Verbeke & Vackier, 2004). There has been a great deal of interest over recent years in the protective effect of fruit and vegetables against a number of diseases, and there is convincing evidence that high intakes of vegetables and fruit are associated with lower risk of chronic diseases (Colgan et al., 2004; Liu et al., 2000; Sargeant et al., 2001). International and national health organisations (NHMRC, 2002; WHO, 2003) have recommended to increase the consumption of vegetables and fruit as an important health and nutrition priority. In the Health Education Authority's Health and Lifestyle Survey of 1993 it was found that the main demographic characteristics that distinguished between low and high fruit and vegetable consumers were age, gender and smoking status (Thompson et al. 1999). These demographic characteristics perhaps result in the strongest variations in intakes of fruit and vegetables, with women reporting higher preference for eating vegetables than men (Thompson et al., 1999; Wardle et al., 2004). However, men reported to like fruit slightly more than women and there was no significant gender difference in attitudes towards fruit and vegetables, although women's attitudes were slightly more positive (Wardle et al., 2004).

One possible mechanism for the gender-specific patterns of healthy food choices might be related to nutritional knowledge. A number of studies have reported gender differences in the knowledge nutritional information (Crawford & Baghurst, 1990; Parmenter et al., 2000; Tate & Cade, 1990), supporting the hypothesis that differences in awareness could contribute to gender differences in intake. Gender, level of education and occupational social class were found to have significant independent effect on level of nutrition knowledge scores. In particular, women demonstrated superior knowledge regarding all the areas of nutrition, as confirmed by the majority of studies dealing with the evaluation of nutritional knowledge (Butriss, 1997; Parmenter et al., 2000). Food-related activities, such as shopping, cooking and eating are conventionally presented as female-centered (Caplan et al., 1998; Warde & Hetherington, 1994). Given women's traditional role in purchasing, preparing and providing food, it is not surprising that men know less about the health benefits of specific food items (Nutrition Forum, UK, 2003). The rise in the number of people living alone together with the decline in the number of traditional family units, where the husband earns and the wife is responsible for shopping and cooking, has raised new concerns. In fact, it appears that even though there is an increasing number of men cooking for themselves and fewer relying on women to make decisions about their diets, this novel activity is not accompanied by a significant increase in nutrition knowledge. However, the significance of nutrition knowledge as an determinant in food choices has been questioned in the light of evidence from research in the field of fat and fiber intake showing no more than small correlations between nutrition knowledge and dietary quality (Shepherd and Towler, 1992; Lappalainen et al., 1997). However, a recent study found substantial associations between knowledge and fruit and vegetable intake, possibly because, unlike the situation for fat intake, overall levels of public awareness are low (Wardle et al., 2000). Fewer men than women knew the current recommendations for fruit and vegetable intake, and fewer were aware of the links between fruit and vegetable consumption and disease prevention. The evidence that men have a lower knowledge about nutrition, or accord lower priority to nutrition in making their food choices, could result in lower intakes of fruit and vegetables. However, only in four member states (Austria, Belgium, Finland and Italy) within the Pan-Europe survey, more men than women reported lack of knowledge as a barrier towards healthy eating (Lappalainen et al., 1997).

Even though methodological differences in assessing food choices might have been in part generated slightly different results on gender-specific food choice, all the observations reported here are consistent in concluding that women generally make slightly healthier food choices. If women report healthier practices (or at least attempt to make healthier choices) all over the world, then this would suggest that any explanations for the differences are more likely to indicate underlying behavioural characteristics of men and women than local cultural effects.

2.2 Eating behaviour and weight control

A factor that could contribute to gender differences in food choice is women's greater concern about weight control and their higher frequency of dieting. There is a consistent body of recent literature (Afifi et al., 2002; Johnson & Wardle, 2005; Kostanski et al., 2004; Liebman et al, 2001; Wardle et al., 2000; Wardle & Griffith, 2001) that clearly indicates that there are important gender differences in weight concern and body self-perception.

Weight control/body perception are known to influence food choice decisions, mainly in women (Glanz et al., 1998; Goode et al., 1995; Rozin et al., 1999). In many studies of attitudes to body weight or dieting, women reported more dissatisfaction with their weight and make more attempts to control weight than men (Beardsworth et al., 2002; Bellisle et al., 1995; Wardle & Griffith, 2001;). Numerous research on body image have shown that women are more likely than men to perceive themselves as overweight and to express discontentment with their body shape (McElhone et al., 1999; Neumark-Sztaine et al., 1999).

Concern with adhering to a slimming diet has been found to be significantly more widespread among women than men (Germov & Williams, 1996; Sobal et al., 1995). A Pan-EU survey on 15239 subjects (European Communities, 1999), reported a strong gender difference in the percentage of people who are content with their body weight. A consistent majority of males was comfortable with their current body weight compared with females. Conversely, a far higher proportion of females wished to be lighter or considerably lighter compared with males. The relative proportion choosing dieting as the strategy for losing weight compared with other methods was highest in the group wishing to be considerably lighter, especially among females (European Communities, 1999). On the other hand, it has been observed that men generally prefer to select physical exercise than dieting as a means for body weight control, while women were more inclined to select dieting, restrained eating and daily checking of body weight (Clark et al., 2009).

Moreover, in women, the frequency of dieting is often associated with difficulties in eating behaviour. Restrained eating behaviour, cognitive control and eating disorders are mainly seen as behavioural phenomena more common in women. Men, on the other hand, have fewer problems with their eating behaviour, and their attitude to food is generally uncomplicated and enjoyable, even though they are more frequently overweight and have higher risk of associated disease (Kiefer et al., 2005).

Problems with eating behaviour have a strong female prevalence emerging in childhood and adolescence(Afifi-Soweid et al., 2002). Girls often eat less and pay attention to calories, sugar and fat intake under the pressure of "feeling obliged" to be slim. Consequently, in part due to a specific social pressure, girls are more likely than boys to develop eating disorders (i.e., anorexia, bulimia, binge eating disorder). Women affected by certain eating disorders are likely to experience a constant internal conflict between the desire of being slim or slimmer, and the drive for certain "forbidden" food. Women are more often affected by the problem of craving (i.e., the strong willing for certain foods) than man, being more likely to be wishful for sweet foods. This attitude results in a difficulty in sticking to a weight reducing-slimming diet (Lafay, et al, 2001). Extensive research showed that women often experience the so called "carbohydrate craving" and there is an association between the wish for sugar- and fat-rich foods (like chocolate and other sweets) and menstrual cycle (Bruinsma and Taren, 1999, Rozin et al., 1991, Smith & Souter, 1969; Yen at al., 2010). Recent findings showed that the wish for sweet food regresses in women with increasing age (Kiefer et al., 2005).

In connection with the craving for particular foods, women more frequently report negative feelings, in contrast to men who describe positive feelings (Lafay et al., 2001). Women also eat more than usual in stressful situations more frequently than men (Kandiah et al., 2006).

A study dealing with the emotional triggers of "comfort" food consumption indicate that in women this eating behaviour was triggered by negative feelings, whereas in men was motivated by positive emotions (Dubè et al., 2005).

Differences in preference towards "comfort" foods across gender were investigated in a survey conducted in North America on 1416 people (Wansink et al., 2003). The findings of this study are consistent with other research showing that females preferred "comfort" foods within the category of snacks, such as chocolate, candy and ice cream. Indeed, one research on "chocolate addiction" reported that 70 of the 72 self-selected "addicts" were female (Tuomisto et al., 1999) and in another study the 92% of the surveyed sample self-identified as "chocolate addicts" were female (Hetherington & MacDiarmid, 1993). Wansink et al. (2003) reported that males preferred more substantial, warm, hearty- meal related comfort foods such as meat dishes, pizza or pasta, casserole and soup. On the other hand side it emerged that when women indulged in high-calories sweets like candy or ice cream often felt guilty afterwards – while men who chose foods other than sweets and snacks, didn't (Wansink et al., 2003).

3. Gender differences in nutrients intake

In order to detect any gender-associated trend in nutrients intake, we have analysed the nutritional profiles estimated at population level within the context of nationwide individual dietary surveys. We considered the database of dietary intakes of 22 European Countries, partners of the European and Health Report (ENHR II), the most suitable source of data on the basis of the number of countries involved and the approach utilized to collect the indicators (Elmadfa, 2009). Data from this report have been therefore processed to highlight possible differences between males and females population groups concerning the percentage contribution to the average daily energy intake by carbohydrates, proteins, fats, fatty acids, and mean daily intake of minerals and vitamins. Data were then grouped by nutrients, by gender, country, and age-class and graphically plotted in order to detect similarities and differences associated to gender. Overall, no significant differences were observed in the percentage contributions of macronutrients to the daily energy intake. Minimum and maximum values (ranges) observed in average per-capita daily intakes in females per each country overlap those of males. These findings are in agreement with data obtained in previous studies (Flynn et al., 2009 ; Reynolds at al., 1999). In the European Health and Nutrition report (ENHR II, Elmafda, 2009), solely for carbohydrates were found differences in the percentage of the average daily energy intake equal or higher than 5% among adults (10% in Estonia, and Lithuania; 9% in Czech Republic; 8% in Austria, Poland, Denmark, and Germany; 7% in Hungary, Portugal, and Finland; 5% in Latvia; less than 5% resulted in Greece, Sweden, France, Spain, The Netherlands, United Kingdom, Romania, Norway, and Italy), and elderly (10% United Kingdom and Denmark; 9% Germany; 7% in Hungary and Romania; 5% Poland and Greece; less than 5% France, Finland, The Netherlands, Sweden, Ireland, Spain, and Italy). The gap is due to the fact that the group of adult and elderly males tends to replace carbohydrates with alcohol. In two cases, differences higher than 10% were found for fats among adolescents in Norway (11% males vs. females) and Dutch (19% females vs. males). Percentage of energy from proteins did not show differences in absolute value higher than 1%.

Mineral	Gender	Age class											
		4-6 years		7-9 years		10-14 years		15-17 years		18-64 years		65+ years	
		min	max	min	max	min	max	min	max	min	Max	min	max
Sodium (g)	Male	1,8	3,4	2,2	4,2	2,3	5,0	2,9	6,5	2,6	7,3	2,3	7,0
	Female	1,7	3,4	1,9	3,7	2,2	4,8	2,2	4,5	1,7	5,6	1,8	6,0
	Δ%	6	0	16	14	5	4	32	44	53	30	28	17
Potassium (g)	Male	2,0	2,8	1,9	3,0	1,9	4,0	2,4	4,4	2,7	4,4	2,2	3,8
	Female	1,9	2,7	1,8	2,9	1,7	4,0	1,2	3,3	2,3	3,6	2,2	3,7
	Δ%	5	4	6	3	12	0	100	33	17	22	0	3
Calcium (mg)	Male	604	1103	732	1207	701	1381	806	1447	687	1171	627	1071
	Female	606	1024	631	1126	600	1238	645	1040	508	1047	533	959
	Δ%	0	8	16	7	17	12	25	39	35	12	18	12
Phosphorus (mg)	Male	882	1284	960	1455	964	1704	1413	1705	1264	1778	1059	1576
	Female	900	1183	851	1295	807	1636	962	1356	1017	1422	959	1404
	Δ%	-2	9	13	12	19	4	47	26	24	25	10	12
Magnesium (mg)	Male	171	267	204	303	200	503	260	467	256	465	221	403
	Female	166	267	166	301	181	429	186	369	197	372	179	348
	Δ%	3	0	23	1	10	17	40	27	30	25	23	16
Iron (mg)	Male	7,3	10,6	8,4	11,8	9,2	19,4	10,2	19	10,6	26,9	10	25,2
	Female	6,8	10,6	7,7	11,8	7,7	14,8	7,8	14	8,2	22,2	8,5	20,9
	Δ%	7	0	9	0	19	31	31	36	29	21	18	21
Zinc (mg)	Male	6,0	9,2	7,0	10,9	7,0	14,6	9,3	15,2	8,6	14,6	7,5	12,3
	Female	5,3	8,9	6,4	9,4	6,1	13,9	6,4	11,0	6,7	10,7	6,7	11,2
	Δ%	13	3	9	16	15	5	45	38	28	36	12	10
Iodine (mg)	Male	69	223	73	203	102	209	93	221	67	264	107	226
	Female	70	198	73	179	85	171	115	182	48	200	97	190
	Δ%	-1	13	0	13	20	22	-19	21	40	32	10	19
Copper (mg)	Male	0,7	2,2	0,9	2,8	0,9	2,9	1,2	3,4	1,1	2,3	1,1	1,9
	Female	0,7	2,0	0,8	2,6	0,7	2,8	0,8	2,1	0,1	2,2	0,9	1,9
	Δ%	0	10	13	8	29	4	50	62	1000	5	22	0
Manganese (mg)	Male	1,4	3,3	1,7	3,7	1,8	4,6	2,5	6,1	2,0	5,0	2,2	4,9
	Female	1,3	3,0	2,0	3,0	1,7	4,1	1,9	4,4	2,2	5,0	2,0	5,0
	Δ%	8	10	-15	23	6	12	32	39	-9	0	10	-2
Selenium (µg)	Male	23	61	27	41	29	110	39	59	36	73	39	62
	Female	24	61	20	50	28	104	30	38	31	54	34	55
	Δ%	-4	0	35	-18	4	6	30	55	16	35	15	13

Table 1. Minerals ranges of mean daily intake by age-class, and gender calculated starting from values observed in 22 European Countries (ENHR II partners).

		Age class											
Vitamin	**Gender**	4-6 years		7-9 years		10-14 years		15-17 years		18-64 years		65+ years	
		min	max	min	max	min	max	min	max	min	max	min	max
Vitamin A-(g)	Male	0,4	1,1	0,4	1,3	0,4	2,4	0,4	1,8	0,5	2,2	0,5	2,5
	Female	0,4	1,2	0,4	1,1	0,3	2,3	0,3	1,6	0,5	2,0	0,4	2,3
	Δ%	0	-8	0	18	33	4	33	13	0	10	25	9
β-carotene (g)	Male	1,2	3,8	1,6	4,1	1,1	4,8	1,2	4,7	1,4	5,3	1,3	4,8
	Female	1,1	3,4	1,6	4,0	1,1	5,2	1,0	4,7	1,4	5,6	1,3	5,0
	Δ%	9	12	0	2	0	-8	20	0	0	-5	0	-4
Vitamin D (µg)	Male	1,8	5,8	1,5	6,4	1,5	4,8	1,8	7,5	1,6	10,9	0,7	15,0
	Female	1,5	6,5	1,5	5,1	1,2	4,5	1,5	7,1	1,2	10,1	0,7	12,9
	Δ%	20	-11	0	25	25	7	20	6	33	8	0	16
Vitamin E (mg)	Male	5,3	9,8	6,3	11,2	5,9	14,5	6,8	20,8	3,3	17,4	6,3	13,7
	Female	5,1	9,8	5,9	13,3	5,6	18,1	6,0	15,5	4,2	16,1	6,7	13,7
	Δ%	4	0	7	-16	5	-20	13	34	-21	8	-6	0
Thiamin (mg)	Male	0,8	1,4	0,9	1,6	0,9	2,3	1,3	2,3	1,1	2,3	0,9	2,1
	Female	0,8	1,3	0,8	1,4	0,8	1,9	1,0	1,9	0,9	2,1	0,9	1,4
	Δ%	0	8	13	14	13	21	30	21	22	10	0	50
Riboflavin (mg)	Male	1,3	2,1	1,2	2,0	1,2	2,9	1,5	2,6	1,4	2,4	1,2	3,2
	Female	1,2	1,8	1,1	1,7	1,1	2,8	1,2	2,3	1,2	2,8	1,2	2,6
	Δ%	8	17	9	18	9	4	25	13	17	-14	0	23
Niacin (mg)	Male	15,7	24,9	18,7	29,9	8,7	40,4	12,2	43,3	9,2	41,3	8	37,9
	Female	14,4	24,6	16,2	26,3	6,9	32,5	7,3	30,5	6,4	30,6	6,7	31,9
	Δ%	9	1	15	14	26	24	67	42	44	35	19	19
Vitamin B6 (mg)	Male	1,3	1,8	1,2	2,5	1,2	2,8	1,5	3,1	1,6	3,5	1,2	3,0
	Female	1,0	1,9	1,1	1,9	1,1	2,7	1,2	2,5	1,3	2,1	1,2	2,9
	Δ%	30	-5	9	32	9	4	25	24	23	67	0	3
Folates (µg)	Male	150	256	144	290	149	428	190	365	203	494	139	343
	Female	109	199	133	264	140	360	154	298	131	392	121	335
	Δ%	38	29	8	10	6	19	23	22	55	26	15	2
Cobalamin (µg)	Male	2,7	5,3	3,6	5,5	3,2	11,8	4,9	7,5	1,9	9,3	3,1	8,2
	Female	2,6	5,0	2,2	5,3	2,2	11,1	3,5	5,2	1,0	8,8	2,5	7,5
	Δ%	4	6	64	4	45	6	40	44	90	6	24	9
Vitamin C (mg)	Male	60	170	63	172	73	197	71	201	64	153	59	142
	Female	61	157	57	172	77	222	67	205	62	153	60	160
	Δ%	-2	8	11	0	-5	-11	6	-2	3	0	-2	-11

Table 2. Vitamins ranges of mean daily intake by age-class, and gender calculated starting from values observed in 22 European Countries (ENHR II partners).

To complete the analysis ranges of average daily intake of minerals and vitamins intake across the 22 European countries were compared (table 1-2), confirming males having higher intakes in general, in almost all age-gender groups looking at both the minimum and the maximum value of intake observed in the 22 countries. Exceptions for minerals (Table 1): calcium and phosphorus (minimum in 4-6 years old), iodine (minimum in 4-6 and 15-17 years old), manganese (minimum 7-9 years old, 18-64 and maximum in 65+ years old), and selenium (minimum in 4-6 years old, maximum in 15-17 years old) (Elmadfa, 2009).

Vitamins intake showed higher variability than minerals intake (Table 2). Males had higher or equal per-capita average daily intakes both for minimum and maximum in all age-gender groups for niacin, folate, cobalamine, and thiamine only. In all other cases at least one age group showed either the minimum or the maximum value

The similarities in the percentage of energy provided by macronutrients and the overlapping of ranges for minerals and vitamins evidenced above indicate that the overall quality of diet does not differ substantially between males and females in all classes of ages and throughout Europe. A further remarks concerns dietary fibre, being highly associated with the mean energy intake, then increasing with the age (Elmadfa, 2009).

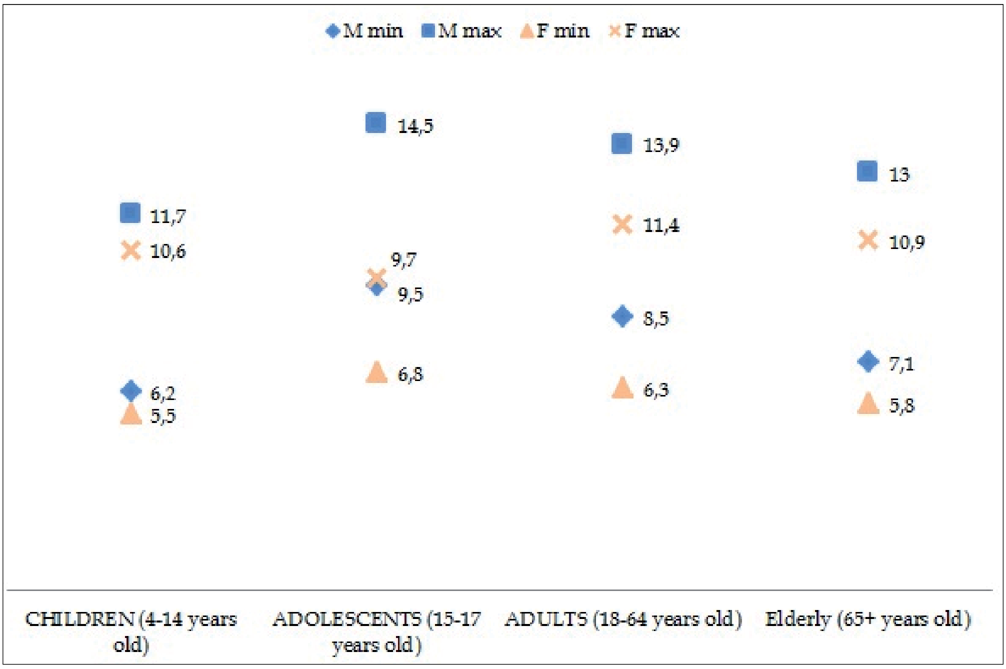


Fig. 1. Ranges of per-capita average daily intake (MJ) by age- class and gender

Overall, the level of energy intake was the main difference found when comparing the average per-capita dietary intakes of males vs. females (Elmadfa, 2009). As expected, men consume more energy than women (Figure 1), similarly to what is reported in other studies like a cross-country comparison (Flynn et al. 2009), or a cross-studies analysis (Reynolds et al. 1999). This difference occurs even though men and women show a similar food volume consumption (Marti-Henneberg et al. 1999).

In other words, the quantity of food does not seem the main component concurring to the energy intake. The combination of food categories characterizing the diet according to gender therefore plays a central role in determining the amount of energy consumed.

3.1 Dietary profiles and nutrient intakes

In order to explore more in depth the relation between food consumption and nutrients intake, we investigated the available literature reporting studies addressing gender specific dietary profiles. According to Kiefer and co-authors (Kiefer et al., 2005) children, adolescents and adults males consume more energy, fat, and cholesterol but less carbohydrates and fibre than females. Fibre intake was found higher in females also in most studies reported by Reynolds et al. (1999). Data from The National Health and Nutrition Examination Survey (NHANES II) on the US population from 1976-1980, indicate that males consume more calories and fats than females (Block et al., 1988).

A research conducted in the US (Courtenay, 2000) showed that males of all ages consume more saturated fat and dietary cholesterol than females. Cholesterol intake of males was substantially higher that recommended levels, while dietary cholesterol of most females of all ages fell within the recommended range for classes of age (Courtenay, 2000).

One study (Wardle et al., 2004) showed that gender differences in food consumption do not always reflect differences in the proportion of energy consumed as fats or fibre intake, but this might be due to gender difference in alcohol consumption, which is likely to add a substantial amount of energy as "drink calories". Once adjusted for energy intake, the dietary micronutrients profile of women tend to be higher than in men. In general, the diets of females were more nutrient-dense, with the exception of milk-derived calcium, and also higher in dietary fibre, phytochemicals, and various micronutrients (Liebman et al., 2003).

Among school children, girls were found to consume much less energy than boys and also have a reduced micronutrient intakes. Pre-school children did not show significant gender differences in dietary profiles (Backstrand et al., 1997). These observations corroborate the importance of differences associated to gender in food choices in determining the quality of the diet, at individual level.

According to Chung & Hoerr (2005) , women seem more predisposed to meet the minimum recommended number of servings of fruit. Moreover, men have been shown to consume less carotenoid-rich foods, such as carrots, spinach, broccoli and other greens than women (Courtenay, 2000).

In agreement with the study by Chung & Hoerr (2005) and Wardle and co-workers (2004) showed that women eat more fruit than men. Several studies have reported that in various western countries women eat more fruits, vegetables, cereals, cereal products, milk, dairy products and whole grain products than men. On the other hand, the consumption of red meat, eggs, alcohol, soft drinks, high sucrose food, as well as various high starch foods such as potatoes and bread is higher in men (Beer-Borst et al., 2000; Fraser et al., 2000; Kiefer, 2005; Prättälä et al., 2007; Wardle, 2004).

The differences associated with gender were similar in all countries, throughout age and educational groups, and in rural and urban areas (Prättälä et al., 2007). A study on gender differences in food intake conducted on1556 older people living in Britain (Fraser et al., 2000), also indicated that women eat more butter, full-fat milk and certain beverages, cakes, apples, pears and bananas, whereas men consume more eggs, sugar and meat products. Men also consume more alcoholic drinks, more frequently and in higher amounts than women (Bates et al., 1999; Fraser et al., 2000;Liebman et al., 2003). Similar trends have been found in two British national surveys, the Health and Lifestyle Survey (Cox et al., 1993), and the National Dietary and Nutritional Survey (Gregory et al., 1990) . These observations are also in agreement with data obtained in studies conducted in the United States (Wirfalt & Jeffrey, 1997) and Australia (Baghurst et al., 1994). On the whole, the studies reported here strongly suggest that a higher intake of fruit and vegetables is one of the elements characterizing women's dietary profiles.

The results of the present analysis on gender differences in nutrients and foods intake lead to figure out that fruit and vegetables consumption is a suitable indicator for dietary characterization. This remark suggests to include this relatively novel quantitative variable in future population studies as a tool to analyse gender specific eating behaviour. Moreover, a composite methodological design linking food choice and dietary intake approaches could help to deepen the knowledge of eating behaviour in the population.

Gender is usually considered in the Nutritional recommendations published by European countries (Pavlovic et al., 2007), even though a conceptual bias, might occur when scientific evidences are collected on males and then extrapolated to women (Tarnopolsky, 2003). Nutritional recommendations at international level are developed by gender and age when addressing specific nutrient requirements (World Health Organization [WHO] 2010a, thereafter (WHO, 2010a) whereas this does not occur in the formulation of nutritional goals for the general population (WHO 2010a, 2010b, 2010c). Further investigations should be carried out to analyse the appropriateness of including gender specific statements in internationals nutrition policy guidelines.

4. Conclusions

All reported data are consistent with the view that in modern Western Societies women generally show a tendency to perform healthier food choices and are much more concerned about the importance of food choice and eating behaviour to stay in a good physical shape than men. This attitude is also reflected by dietary profiles in terms food intake pattern, showed consistent trends according to gender.

In conclusion, the "take home message" extrapolated from this systematic review, in strong accordance with other recent studies (Berbesque, 2009; Marino et al., 2011), is to stress the importance of considering a gender specific approach, both in terms of behaviour and of physiology, when addressing nutrition issues in research and in policy making. As a matter of fact, a more detailed informative basis would help the formulation of suitable monitoring programs in the research side, and an increasing effectiveness of policy interventions in respect of different population groups.

Acknowledgment. The present review was performed within the project PALINGENIO supported by the Italian Ministry of Agriculture, Food and Forestry Policy.

5. References

Afifi-Soweid, R.A., Kteily, M.B.N., & Shediac-Rizkallah, M.C. (2002). Preoccupation with weight and disordered eating behaviours of entering students at a university in Lebanon. *International Journal of Eating Disorders*, Vol. 32, No. 1, pp. 52-57

Axelson, M. L., & Brinberg, D. (1989). *A social-psychological perspective on food-related behaviour*. Springer-Verlag, New York

Backstrand, J.R., Allen, L.H., Pelto, J.H., & Chàvez, A. (1997). Examining the gender gap in nutrition an example from rural Mexico. *Social Science & Medicine*, Vol. 44, pp. 1751-1759

Baghurst, K.I., Baghurst, P.A., Record, S.J. (1994). Demographic and dietary profiles of high and low fat consumers in Australia. *Journal of Epidemiology & Community Health*, Vol. 48, pp. 26-32

Baker, A. H., & Wardle, J. (2003). Sex differences in fruit and vegetable intake in older adults. *Appetite*, 40(3), 269-275.

Barker, M.E., Thompson, K.A., & McClean, S.I. (1995). Attitudinal dimensions of food choice and nutrient intake. *British Journal of Nutrition*, Vol. 74, pp. 649–659

Bates, C.J., Prentice, A., & Finch, S. (1999). Gender differences in food and nutrient intakes and status indices from the National Diet and Nutrition Survey of people aged 65 years and over. *European Journal of Clinical Nutrition*, Vol. 53, pp. 694-699

Beardsworth, A., Bryman, A., Keil, T., Goode, J., Haslam, C. & Lancashire, E. (2002). Women, men and food: the significance of gender for nutritional attitudes and choices. *British Food Journal*, Vol. 104, pp. 470-491

Beer-Borst, S., Hercberg, S., Morabia, A., Bernstein, M.S., Galan, P., Galasso, R., Giampaoli, S., McCrum, E., Panico, S., Preziosi, P., Ribas, L., Serra-Majem, L., Vescio, M.F., Vitek, O., Yarnell, J., Northridge, M.E. (2000). Dietary patterns in six European populations: results from EURALIM, a collaborative European data harmonization and information campaign. *European Journal of Clinical Nutrition*, Vol. 54, pp. 253–262

Bellisle, F., Monneuse, M. O., Steptoe, A., & Wardle, J. (1995). Weight concerns and eating patterns: A survey of university students in Europe. *International Journal of Obesity and Related Metabolic Disorders*, Vol. 19, pp. 723–730

Berbesque, J. (2009). Sex differences in food preference of Hazda Hunter-Gatherers. *Evolutionary Psychology*, Vol. 7, No. 4, pp. 601-616

Block, G., Rosenberger, W.F., Patterson, B.H. (1988). Calories, fat and cholesterol: intake patterns in the US population by race, sex and age. *American Journal of Public Health*, Vol. 78, pp. 1150-1155

Booth, D.A., & Shepherd, R. (1988). Sensory influences on food acceptance: the neglected approach to nutrition promotion. *British Nutrition Foundation Nutrition Bulletin*, Vol. 13, pp. 39-54

Booth, D. (1994). *The psychology of nutrition*. Taylor & Francis, Bristol, UK

Brendan, G. (2007). 'Real men don't diet': An analysis of contemporary newspaper representations of men, food and health, *Social Science & Medicine*, Vol. 64, pp. 326–337

Bruinsma, K., Taren, D.L. (1999). Chocolate: Food or Drug? *Journal of the American Dietetic Association*, Vol. 99, No. 10, pp. 1249-1256

Buttriss, J.L. (1997). Food and nutrition: Attitudes, beliefs and knowledge in the United Kingdom. *American Journal of Clinical Nutrition*, 65, 1985S–1995S

Caplan, P., Keane, A., Willetts, A., & Williams, J. (1998). Concepts of healthy eating: Approaches from a social science perspective. In A. Murcott (Ed.), *The nation's diet: The social science of food choice.* Longman, London

Chung S., & Hoerr S. (2005). Predictors of fruit and vegetable intakes in young adults by gender. *Nutrition Research*, Vol. 25, pp. 453-463

Clarke, P.J., O'Malley, P.M., Johnston, L.D., Schulenberg, J.E., & Lantz, P. (2009). Differential Trends in Weight-Related Health Behaviours Among American Young Adults by Gender, Race/Ethnicity, and Socioeconomic Status: 1984-2006. *American Journal of Public Health*, Vol. 99, No. 10, pp. 1893-1901

Colgan, H.A., Floyd, S., Noone, E.J., Gibney, M.J., & Roche, H.M. (2004). Increased intake of fruit and vegetables and a low-fat diet, with and without low-fat plant sterol enriched spread consumption: effects on plasma lipoprotein and carotenoid metabolism. *Journal of Human Nutrition and Dietetics*, 17, pp. 561–569

Conner, M., & Armitage, C.J. (2002). *The social psychology of food*. Open University Press, Buckingham, United Kingdom.

Counihan, C. M. (1999). *The anthropology of food and body*. Routledge, New York

Courtenay, W.H., (1998). College men's health: an overview and a call to action. *Journal of American College Health*, Vol. 46, No. 6, pp. 279-290

Courtenay, W.H. (2000). Behavioural factors associated with disease, injury, and death among men: evidence and implications for prevention. *Journal of Men's Studies*, Vol. 9, No. 1, pp. 81-142

Cox, B.D., Huppert, F.A., & Whichelow, M.J. (1993). The health and lifestyle survey: seven years on. Aldershot, United Kingdom

Crawford, D.A., & Baghurst, K.I. (1990). Diet and health: a national survey of beliefs, behaviours and barriers to change in the community. *Australian Journal of Nutrition and Dietetics*, Vol. 47, pp. 97–104

Dubé, L., LeBelm, J.L., & Lu, J. (2005). Affect asymmetry and comfort food consumption. *Physiology & behaviour*, Vol. 86, No. 4, pp. 559-567

Elmadfa, I. (Ed.) (2009). European Nutrition and Health Report 2009. *Annals of Nutrition and Metabolism* 55(2) pp. I-V. Available from http://www-univie.ac.at/enhr/

European Communities (1999). A pan-EU survey on consumer attitudes to physical activity, body weight and health. Employment & Social Affairs-Health, Directorate V/F.3 ISBN 92-828-5421-3. Office for Official Publications of the European Communities. L-2985 Luxemburg

EUROSTAT – Eurostat, the statistical office of the European Union (2011). Statistics Database – Data Navigation Tree: *tsiir050*, Available at http://epp.eurostat.ec.europa.eu/tgm/refreshTableAction.do?tab=table&plugin=1&pcode=tsiir050&language=en

Fagerli, R.A., & Wandel, M. (1999). Gender Differences in Opinions and Practices with Regard to a "Healthy Diet". *Appetite*, Vol. 32, pp. 171–190

Fischler, C. (1980). Food Habits, Social Change, and the Nature/culture Dilemma. *Social Science Information*, Vol. 19. No. 6, pp. 937-953

Flynn, A., Hirvonen, T., Mensink, G.B., Ocké, M.C., Serra-Majem, L., Stos, K., Szponar, L., Tetens, I., Turrini, A., Fletcher, R., & Wildemann, T. (2009). Intake of selected nutrients from foods, from fortification and from supplements in various European countries, *Food and Nutrition Research*, suppl. 1, pp. 1-51 http://www.foodandnutritionresearch.net/index.php/fnr

Fraser, G.E., Welch, A., Luben, R., Bingham, S.A, & Day, N.E. (2000). The effect of age, sex, and education on food consumption of a middle-aged English cohort-EPIC in East Anglia. Preventive Medicine, Vol. 30, pp. 26-34

Friel, S., NicGabhainn, S., & Kelleher, C. (1999). Main results of the national health and lifestyle surveys, SLAN and HBSC. Centre for Health Promotion Studies: National University of Ireland and Department of Health and Children, Galway

Fulkerson, J., French, S., & Story, M. (2004). Adolescents' attitudes about and consumption of low-fat foods: associations with sex and weight-control behaviours. Journal of the American Dietetic Association, Vol. 104, No. 2, pp. 233-237

Germov, J., & Williams, L. (1996). Sexual division of dieting: women's voices. *Sociological Review*, Vol. 44, No. 4, pp. 630-647

Girois, S.B., Kumanykia, S.K., Morabia, A., & Mauger, E. (2001). A comparison of knowledge and attitudes about diet and health among 35 to 75 year-old adults in the United States and Geneva, Switzerland. *American Journal of Public Health*, Vol. 91, pp. 418–424

Glanz, K., Hewitt, A.M., & Rudd, J. (1992). Consumer behaviour and nutrition education: an integrative review. *Journal of Nutrition Education,* Vol. *24,* pp. 267–277

Glanz, K., Basil, M., Maibach, E., Goldberg, J., & Snyder, D. (1998). Why Americans eat what they do: taste, nutrition, cost, convenience, and weight control concerns as influences on food consumption. *Journal of the American Dietetic Association*, Vol. 98, No. 10, pp. 1118–1126

Goode, J., Beardsworth, A., Haslam, C., Keil, T., & Sherratt, E. (1995). Dietary dilemmas: nutritional concerns of the 1990's. *British Food Journal*, Vol. 11, pp. 3–12

Gough, B., & Conner, M. T. (2006). Barriers to healthy eating amongst men: A qualitative analysis. Social Science & Medicine, Vol. 62, No. 2, pp. 387–395

Gregory, J., Foster, K., Tyler, H., & Wiseman, M. (1990). *The Dietary and Nutritional Survey of British Adults*. The Stationery Office, London

Harris, M. (1985). *Good to Eat: Riddles of Food and Culture* (Previously published as *The Sacred Cow and the Abominable Pig*). 2nd Edition 1998, Waveland Press Inc., Long Grove, Illinois USA

Hearty, A.P., McCarthy, S.N., Kearney, J.M., & Gibney, M. J. (2007). Relationship between attitudes toward healthy eating and dietary behaviour, lifestyle and demographic factors in a representative sample of Irish adults. *Appetite*, 48, 1-11

Hetherington, M.M., & MacDiarmid, J.I. (1993). Chocolate addiction: a preliminary study of its description and its relationship to problem eating. *Appetite*, Vol. 21, pp. 233–246

Herman, C.P., Roth, D. A., & Polivy, J. (2003). Effects of the presence of others on food intake: A normative interpretation. *Psychological Bulletin*, Vol. 129, pp. 873–886

Hunt, M.K., Stoddard, A.M., Glanz, K., Hebert, J.R., Probart, C., Sorensen, G., Thomson S., Hixson M.L., Llnnan, L., Palombo, R., For The Working Well Research Group (1997). Measures of Food Choice Behaviour Related to Intervention Messages in Worksite Health Promotion. *Journal of Nutrition Education*, Vol. 29, No. 1, pp. 3-11

Jensen, K.O., & Holm, L. (1999). Preferences, quantities and concerns: Sociocultural perspectives on the gendered consumption of foods. *European Journal of Clinical Nutrition*, Vol. 53, No. 5, pp. 351–359

Johnson, F., & Wardle, J. (2005). Dietary restraint, body dissatisfaction, and psychological distress: A prospective analysis. *Journal of Abnormal Psychology*, Vol. 114, No. 1, pp. 119-125

Kandrack, M., Grant, K.R., & Segall, A. (1991). Gender differences in health related behaviour: some unanswered questions. *Social Science and Medicine*, Vol. 32, No. 5, pp. 579-590

Kearney, J. M., Gibney, M. J., Livingstone, M. B., Robson, P. J., Kiely, M., & Harrington, K. E. (2001). Attitudes towards and beliefs about nutrition and health among a random sample of adults in the Republic of Ireland and Northern Ireland. *Public Health Nutrition*, Vol. 4, pp. 1117–1126

Kiefer, I., Rathmanner, T., & Kunze, M. (2005). Eating and dieting differences in men and women. *Review Literature And Arts Of The Americas*, Vol. 2, pp. 194-201

Kostanski, M., Fisher, A., & Gullone, E. (2004). Current conceptualization of body image dissatisfaction: Have we got to wrong? *Journal of child Psychology and Psychiatry*, Vol. 45, No. 7, pp. 1317-1325

Lafay, L., Thomas, F., Mennen, L., Charles, M.A., Eschwege, E., Borys, J.M., et al. (2001). Gender differences in the relation between food cravings and mood in an adult community: Results from the fleurbaix laventie ville sante study. *International Journal of Eating Disorders*, Vol. 29, pp. 195–204

Lappalainen, R., Saba, A., Holm, L., Mykkanen, H., & Gibney, M.J. (1997). Difficulties in trying to eat healthier: descriptive analysis of perceived barriers for healthy eating. *European Journal of Clinical Nutrition*, Vol. 51, suppl.2, pp. 36-40

Lennernäs, M., Fjellström C., Becker, W., Giachetti, I., Schmitt, A., Remaut de Winter, A.M., and Kearney. M. (1997). Influences on food choice perceived to be important by nationally-representative sample of adults in the European Union. *European Journal of Clinical Nutrition,* 51, suppl.2, pp.8-15

Li, R. Serdula, M., Bland, S., Mokdad, A., Bowman, B., & Nelson, D. (2000). Trends in fruit and vegetable consumption among adults in 16 US states: Behavioural Risk Factor Surveillance System, 1990-1996. *American Journal of Public Health*, Vol. 90, No. 5, pp. 777–781

Liebman, M., Cameron, B.A., Carson, D.K., Brown, D.M., & Meyer, S.S. (2001). Dietary Fat Reduction Behaviours in College Students: Relationship to Dieting Status, Gender and Key Psychosocial Variables, *Appetite*, Vol. 36, pp. 51-56

Liebman, M., Propst, K., Moore, S.A., Pelican, S., Holmes, B., Wardlaw, M.K., et al. (2003). Gender differences in selected dietary intakes and eating behaviours in rural communities in Wyoming, Montana, and Idaho, *Nutrition Research*, Vol. 23, pp. 991-1002

Lindeman, M., & Sirelius, M. (2001). Food choice ideologies: The modern manifestation of normative and humanist views of the world. *Appetite*, Vol. 37, pp. 175–184

Lindeman, M., & Stark, K. (1999). Pleasure, pursuit of health, or negotiation of identity? Personality correlates of food choice motives among young and middle-aged women. *Appetite*, Vol. 33, pp. 141–161

Lindeman, M., & Stark, K. (2000). Loss of pleasure, ideological food choice reasons and eating pathology. *Appetite*, Vol. 35, pp. 263–265

Liu, S., Manson, J.E., Lee, I.M., Cole, S.R., Hennekens C.H., Willett W.C., et al. (2000). Fruit and vegetable intake and risk of cardiovascular disease: the Women's Health Study. *American Journal of Clinical Nutrition*, 72, pp.922–8

Lonnquist, L.E., Weiss, G.L., & Larsen, D.L. (1992). Health value and gender in predicting health protective behaviour. *Women and Health*, Vol. 19. No. 2/3, pp. 69-85

Marino, M., Masella, R., Bulzomi P., Campesi, I., Malorni, W., & Franconi, F. (2011). Nutrition and human health from a sex–gender perspective. *Molecular Aspects of Medicine*. Vol. 32, No. 1, pp. 1-70

Margetts, B.M., Martinez, J.A., Saba, A., Holm, L., Kearney, M. (1997). Definitions of 'healthy' eating: a Pan-EU survey of consumers attitudes to food, nutrition and health. *European Journal of Clinical Nutrition*, Vol. 51, suppl.2, pp. 23-29

Marti-Henneberg, C., Capdevila, F., Arija, V., Cucó, G., Vizmanos, B., & Fernández-Ballart, J. (1999). Energy density of the diet, food volume and energy intake by age and sex in a healthy population. *European Journal of Clinical Nutrition*, Vol. 53, pp. 421-428

McElhone, S., Kearney, J.M., Giachetti, I., Zunft, H.J.F., Martínez, J.A. (1992) Body image perception in relation to recent weight changes and strategies for weight loss in a nationally representative sample in the European Union. *Public Health Nutrition,* 2, pp. 143-151

Mennell, S., Murcott, A. & Otterloo, A. H. (1992). *The sociology of food: eating, diet, and culture*. Sage, London.

Murcott, A. (1983). *The sociology of food and eating*. Aldershot, Hants, England.

Neumark-Sztainer, D., Story, M., Falkner, N., Beuhring. T., & Resnick, M. (1999). Sociodemographic and personal characteristics of adolescents engaged in weight loss and weight/muscle gain behaviours: Who is doing what? *Preventive Medicine*, Vol. 28, pp. 40–50

Parmenter, K., Waller, J. & Wardle, J. (2000). Demographic variation in nutrition knowledge in England, *Health Education Research*, Vol. 15, No. 2, pp. 163-174

Pavlovic, M., Prentice, A., Thorsdottir, I., Wolfram, G, Branca, F. (2007). Challenges in harmonizing energy and nutrient recommendations in Europe, *Annals of Nutrition and Metabolism*, Vol. 51, pp. 108-114

Pollard, J., Kirk, S.F.L., & Cade, J.E. (2002). Factors affecting food choice in relation to fruit and vegetable intake: A review. *Nutrition Research Reviews*, Vol. 15, pp. 373–387

Povey, R., Conner, M., Sparks, P., James, R., & Shepherd, R. (2000). Application of the theory of planned behaviour to two health-related dietary behaviours: Roles of perceived control and self-efficacy. *British Journal of Health Psychology*, 5, pp. 121–139.

Prättälä, R., Paalanen,L., Grinberga, D., Helasoja, V., Kasmel, A. & Petkeviciene, J. (2007). Gender differences in the consumption of meat, fruit and vegetables are similar in Finland and the Baltic countries. *The European Journal of Public Health*, Vol. 17, No. 5, pp. 520-525

Randall, E., & Sanjur, D. (1981). Food preferences: their conceptualisation and relationship to consumption, *Ecology of Food and Nutrition*, Vol. 11, pp. 151-161

Ree, M., Riediger, N., & Moghadasian, M.H. (2008). Factors affecting food selection in Canadian Population, *European Journal of Clinical Nutrition*, Vol. 62, pp. 1255–1262

Reynolds, K.D., Baranowski, T., Bishop, D.B., Farris, R.P., Binkley, D., Nicklas, T.A., & Elmer, P.J. (1999). Patterns in child and adolescent consumption of fruit and vegetables: effects of gender and ethnicity across four sites, *Journal of the American College of Nutrition*, Vol. 18, pp. 248-254

Roininen, K., Tuorila, H., Zandstra, E. H., de Graaf, C., Vehkalahti, K., & Stubenitsky, et al. (2001). Differences in health and taste attitudes and reported behavior among Finnish, Dutch and British consumers: A cross-national validation of the Health and Taste Attitude Scales (HTAS). *Appetite*, 37, pp. 33–45.

Roos, G., Prattala, R., & Koski, K. (2001). Men, masculinity and food: Interviews with Finnish carpenters and engineers, *Appetite*, Vol. 37, No. 1, pp. 47–56

Rozin, P., Levin, E., & Stroess, C. (1991). Chocolate craving and liking. *Appetite*, Vol. 17, pp. 199–212

Rozin, P., Fischler, C., Imada, S., Sarubin, A., & Wrzesniewski, A. (1999). Attitudes to food and the role of food in life in the U.S.A., Japan, Flemish Belgium and France: possible implication for the diet-health debate, *Appetite*, Vol. 33, pp. 163-180

Shepherd, R. (1990). Overview of factors influencing food choice. In: M. Ashwell (Ed.), *Why we eat what we eat*, pp. 12–30. British Nutrition Foundation, London

Shepherd, R., & Towler, G. (1992). Nutrition knowledge, attitudes and fat intake: application of the theory of reasoned action. *Journal of Human, Nutrition and Dietetics,* Vol. 5, pp. 387–397

Shepherd, R. (2005). Influences on food choice and dietary behaviour. In: *Diet Diversification and Health Promotion*, I. Elmadfa (Ed.), pp. 36-43, Karger, Basel

Sobal, J., Nicolopoulos, V., & Lee, J. (1995). Attitudes about overweight and dating among secondary school students. *International Journal of Obesity*, Vol. 19, pp. 376-381

Sobal, J. (2005). Men, meat, marriage: Models of masculinity. *Food & Foodways*, Vol. 13, No. 1/2, pp. 135–158

Smith, S.L., & Sauder, C. (1969). Food cravings, depression, and premenstrual problems. *Psychosomatic medicine*, Vol. 31, No. 4, pp. 281-287

Sparks, P., Conner,M., James, R., Shepherd, R., & Povey, R. (2001). Ambivalence about health-related behaviors: An exploration in the domain of food choice. *British Journal of Health Psychology*, 6,pp. 53–68

Steptoe, A. Wardle, J., Cui W., Bellisle F., Zotti A.M., Baranyai R., & Sanderman R. (2002). Trends in Smoking, Diet, Physical Exercise, and Attitudes toward Health in European University Students from 13 Countries, 1990–2000. *Preventive Medicine*, Vol. 35, No. 2, pp. 97-104

Tarnopolsky, M.A. (2003). Females and males: should nutritional recommendations be gender specific? *Schweizerische Zeitschrift für Sportmedizin und Sporttraumatologie*, Vol. 51, pp. 39-46

Tate, J., & Cade, J. (1990). Public knowledge of dietary fat and coronary heart disease. *Health Education Journal*, Vol. 49, pp. 32–35.

Teratanavat, R., & Hooker, N.H. (2006). Consumer valuations and preference heterogeneity for a novel functional food. *Journal of Food Science*, Vol. 71, pp. S533–S541

Thompson, D.M.H. (1988). *Food acceptability*, Elsevier, New York

Thompson, R.L., Margetts, B.M., Speller, V.M., & McVey, D. (1999). The Health Education Authority's health and lifestyle survey 1993: who are the low fruit and vegetable consumers? *Journal of Epidemiology & Community Health*, Vol. 53, pp. 294–299

Thiele S, Weiss C. (2003). Consumer demand for food diversity: evidence for Germany. *Food Policy*, Vol. 28, pp. 99–115

Tuomisto, T., Hetherington, M.M., Morris, M-F., Tuomisto, M.T., Turjanmaa, V., Lappalainen, R. (1999). Psychological and physiological characteristics of sweet food "addiction". *International Journal of Eating Disorder*, Vol. 25, pp. 169–175

Vartaniana, L.R., Hermanb, C.P., Polivy, J. (2007). Consumption stereotypes and impression management: How you are what you eat. *Appetite*, Vol. 48, pp. 265–277

Verbeke, W., & Vackier, I. (2004). Profile and effects of consumer involvement in fresh meat. *Meat Science,* Vol. 67, pp. 159–168

Wadolowska, L., Babicz-Zielinska, E., & Czarnocinska, J. (2008). Food choice models and their relation with food preferences and eating frequency in the Polish population: POFPRES study, *Food Policy*, Vol. 33, No. 2, pp. 122-134

Wansink, B., Cheney, M., Chan, N. (2003). Exploring comfort food preferences across age and gender, *Physiology & Behaviour*, Vol. 79, No. 4/5, pp. 739-747

Warde, A., & Hetherington, K. (1994). English households and routine food practices: A research note, *The Sociological Review*, Vol. 42, No. 4, pp. 758–778

Wardle, J., Parmenter, K., & Waller, J. (2000). Nutrition knowledge and food intake, *Appetite*, Vol. 34, pp. 269–275

Wardle, J., Griffith, J (2001). Socioeconomic status and weight control practices in British adults, *Journal of Epidemiology & Community Health*, Vol. 55, pp. 185-190.

Wardle, J., Haase, A.M., Steptoe, A., Nillapun, M., Jonwutiwes, K., & Bellisle, F. (2004). Gender differences in food choice: the contribution of health beliefs and dieting. *Annals of Behavioural Medicine*, Vol. 27, No. 2, pp. 107-116

Wirfalt, A.K., & Jeffery, R.W. (1997). Using cluster analysis to examine dietary patterns: nutrient intakes, gender, and weight status differ across food pattern clusters, *Journal of American Dietetic Association*, Vol. 97, pp. 272-279

WHO - World Health Organisation (2003). Food Based Dietary Guidelines in the WHO European region, Available from
http://www.euro.who.int/Document/E79832.pdf

WHO – World Health Organisation (2010a). Dietary recommendations / Nutritional requirements Establishing human nutrient requirements for worldwide application, 2010 May update, Available from
http://www.who.int/nutrition/topics/nutrecomm/en/index.html

WHO – World Health Organisation (2010b). 5. Population nutrient intake goals for preventing diet-related chronic diseases, 2010 May update, Available from
http://www.who.int/nutrition/topics/5_population_nutrient/en/

WHO – World Health Organisation (2010c) Vitamin and Minerals requirements in human Nutrition, 2010 May update, Available from
http://whqlibdoc.who.int/publications/2004/9241546123_annexes.pdf

Yen, J.Y., Chang, S.J., Ko, C.H., Yen, C.F., Chen, C.S., Yeh, Y.C., Chen, C.C. (2010). The high-sweet-fat food craving among women with premenstrual dysphoric disorder: emotional response,implicit attitude and rewards sensitivity. *Psychoneuroendocrinology.* Vol. 35, No. 8, pp. 1203-1212

Yokoo, Y. (1996). The Barriers against Women Majoring in Science and Engineering, NISTEP (National Institute of Science and Technology Policy, Japan). Report n. 44,Available from: http://www.nistep.go.jp/achiev/abs/eng/rep044e/rep044ae.html

5

Iron Food Fortification for the Control of Childhood Anemia in Brazil

Joel Alves Lamounier[1], Flávio Diniz Capanema[2] and Daniela Silva Rocha[3]
[1]*Universidade Federal de São João Del Rei,*
Campus Centro Oeste (UFSJ), Divinópolis, Minas Gerais,
[2]*Children and Adolescents Health Nucleun,*
Faculdade da Saúde e Ecologia Humana (FASEH), Vespasiano, Minas Gerais,
[3]*Universidade Federal da Bahia, Vitoria da Conquista, BA,*
Brazil

1. Introduction

Iron deficiency anemia represents a serious nutritional problem worldwide, and it especially affects children and pregnant women in developing countries. According to Global Health Burden Disease Report of World Health Organization, anemia is considered the most prevalent public health problem in the World (World Health Organization [WHO], 2008), which requires public policies to combat iron deficiency and anemia. In Brazil, the National Survey of Demography and Health of Children and Women - PNDS - showed 20.9% prevalence of anemia among children 6 to 59 months, with the highest prevalence in the northeast (Brazilian Institute of Geography and Statistics [IBGE], 2010). Studies conducted between 1996 and 2007, involving children less than 5 years in different regions of the country, showed very high rates of anemia. The prevalences were 47.8-63.7% in south, 10.4-77.8% in southern, 55.1-84% in north, 35.7-89.1% in northeast, and 31-63.1% in midwest of Brazil (Jordão et al., 2009).

A meta-analysis study which included articles published in the last 10 years, the prevalence of anemia was higher than 40.0% in children under seven years old (Andrade, 2004). Also in day care centers iron deficiency anemia is considered the most common nutritional deficiency in childhood, with high prevalence (Capanema et al., 2008; Castro et al., 2011; Matta et al., 2005; Morais et al., 2005; Neves et al., 2005)

The magnitude of nutritional anemia in Brazil represents a serious public health problem, specially the short and long term affects anemia can have on growth in at-risk groups. Thus, as a result of this problem, various intervention approaches are being adopted in attempts to control and prevent anemia. One possible intervention measure, which has shown to be successful in reducing anemia in at-risk groups, is the iron fortification of foods made available to children.

Food fortification is highlighted as one of the most cost-effective health solutions to fight malnutrition among children and anemia deficiencies among women. Fortification of staple foods improves micronutrient status by delivering small amounts of micronutrients on a daily basis. In addition, some other options, such as drinking water for consumption also has been a good alternative. Staple food fortification is routinely practiced around the globe in developed countries and contributed for decreasing childhood anemia. In developing countries, anemia is still a public health problem in a large number of these. Therefore, food fortification with iron have been considered, since it requires no change in eating habits and delivers benefit through the consumption of fortified staple foods or drinking water.

In this chapter this issue was discussed as a Brazilian experience to control and reduce anemia in childhood using the iron fortification strategy. The use of food vehicles, iron salts and their costs, as well as recent works on iron fortification of foods in Brazil are reviewed.

2. General aspects of iron fortification of infant foods

The fortification of foods consists in the addition of complementary nutrients to foods *in natura*. A concern with nutritional deficiencies in populations and a utilization of fortification as an intervention measure were extensively documented throughout the twentieth century. In 1910, for example, in Denmark, due to concern over vitamin A deficiency, which affected large numbers of children, health officials initiated large scale industrial fortification of margarine with vitamin A, resulting in the elimination of xerophthalmia in the population (Nilson & Piza, 1998).

Once foods are enriched with micronutrients, such as iron, large, at-risk populations will be reached over long periods without the need of effective individual cooperation (Tuma et al., 2003; WHO, 1989). Therefore, food fortification is considered highly effective and flexible, is socially acceptable and furthermore, it does not interfere with the population’s dietary habits. In addition, the risk of side effects and toxicity are minimal due to reduced doses of micronutrients added to foods (Tuma et al., 2003)

Food fortification is a public health measure, and in order to be successful, several considerations should be kept in mind. First, the food vehicle of choice must be consumed regularly and in large scale by the targeted population. In addition, the selected food vehicle should be evaluated for potent absorption inhibitors, and if the added iron compound will have an impact on the iron status of the consumer. Secondly, it is important that the selected iron compound does not cause unacceptable changes in color and flavor when added to foods. Additionally, the food vehicle should be sufficiently stable during long periods of storage and during cooking in order to guarantee that true food consumption may be quantitatively capable of contributing significantly to the nutritional requirements of the population. Finally, the food vehicle must be centrally produced and proper technology is available for industrial-scale fortification (Andrade, 2004; Cardoso & Penteado, 1994).

The objective of iron food fortification programs is to increase the dietary mineral in foods to prevent and control iron-deficiency in at-risk groups (Andrade, 2004). The fortification of foods with iron is a preferred strategy advocated by the World Health Organization. Iron added to foods has been shown to be the most efficient options to control iron-deficiency, and studies have shown improvements over a period of one to three months in people suffering from this deficiency (Nilson & Piza, 1998).

In Europe, some countries have adopted a policy of distribution of infant formula and fortified cereals, which resulted in decreasing the prevalence of iron deficiency in last decades (Hercberg, 2001). In the United States, in a cross-sectional study using data from the Centers for Disease Control and Prevention's (CDC), five American States found that the prevalence of anemia among children dropped by more than 50% in the last two decades and was attributed to better nutritional conditions related to large-scale consumption of fortified foods and possibly better iron bioavailability in some products (Sherry et al., 2001). The prevalence of anemia (NHANES III - conducted between 1988 to 1994) in the U.S. was 3% and 9% in children aged one to two years and less than 1% and 3% for children aged three to five years with anemia iron deficiency and iron deficiency, respectively (Looker, 1997).

In Chile the prevalence of iron deficiency anemia is low in infants, preschoolers, school children, adolescents, adult men and women of childbearing age. Only pregnant women are still highly prevalent. It is likely that this low prevalence is due to fortification of flour with iron and B vitamins. The National Program for Complementary Alimentación (PNAC) distributes milk to children since 1952, while since 1970, this is enriched with iron. Through studies, the composition of output has been modified and now provides the program for infants and pregnant women milk powder fortified with iron, zinc, copper and ascorbic acid (Nilson & Piza, 1998). In Panama, children receive free via Alimentación Complementary Program (CAP) cereal fortified with vitamins and minerals. School-age children receiving iron-fortified milk and biscuits since 2006 that are offered to students in the country have also been fortified with iron and other vitamins and minerals, with coverage in difficult areas (Fontes, 2007). Cuba also adopted as a strategy to combat anemia food fortification, and the flours are enriched with iron and other vitamins and minerals since 1999. Children under two years are a priority for action, and more than 95% of the nation's children receive at subsidized prices, a pope fruit enriched with iron and vitamin C (since 2001). Milk fortified with iron is distributed, also at subsidized prices, to children under one year since 2005, and the program has covered 98% (Herrera, 2007).

For children under one year, the most appropriate strategy seems to be the fortification of child foods at home. Fortification of complementary foods (weaning foods targeted to children age) is an alternative to targeted supplementation. Commercially-prepared complementary foods typically reach higher income, more urbanized households and this tends to have been left more to the market as an initiative. Zlotkin and colleagues at the Hospital for Sick Children, University of Toronto (Canada) developed a less costly alternative of the provision of micronutrients which an be added to infant foods. It was named "home fortification" and used sprinkles, the multiple-micronutrient sachet. The biological efficacy, bioavailability, safety and acceptability of Sprinkles were tested in various scenarios, including countries such as Bangladesh, Benin, Bolivia, China, Canada, Ghana, Guyana, Haiti, India, Indonesia, Kyrgyzstan, Mexico, Mongolia, Pakistan, Vietnam (Sprinkles Global Health Initiative, 2009). Bolivia was the first country that has documented the use of home fortification with intervention at the level of public health. In 2005, the country's data pointed to 70% prevalence of anemia among children 6 to 24 months. The country adopted the strategy of distributing sachets containing iron, vitamin A, vitamin C, folic acid and zinc for all children. Each child receives one sachet per day in one meal. Some recent studies look at the costs and potential impacts of sprinkles and conclude that the benefit: cost ratio of sprinkles interventions, containing iron as well as other micronutrients, can be as high as 37:1 if one assumes that a course of intervention for four months between the ages of 6 months and one year largely protects an infant against anemia throughout childhood (Sharieff et al., 2006).

In Latin and South America, food fortification is widely practiced and can be classified in three program types: mandatory fortification of foods commonly consumed in large part by the population, such as wheat flour and corn meal; fortification targeting specific groups as in the example of foods consumed by infant and children populations, in this case cereal, powdered milk, biscuits and other industrialized products; and voluntary fortification, in which the food industry adds iron and other micronutrients to industrialized foods. The direct costs of food fortification are extraordinarily low when compared to the high social costs of micronutrient deficiencies. In most cases, according to the World Bank, the cost of fortification is less than one dollar per year to protect an individual against vitamin A, iron and iodine deficiencies. The cost to prevent an iron-deficiency alone has been estimated to be less than US$ 0.10 per year (Nilson & Piza, 1998).

3. Wheat and corn flour iron fortification

In Brazil, since 2001 the Ministry of Health made mandatory the addition of iron [30% Recommended Nutritional Intake (RNI) or 4.2 mg/100 g] and folic acid (70% RNI or 150 µg) to milled wheat and corn flour. Federal law now dictates mandatory fortification of iron instead of voluntary fortification by the grain industry. This measure has as its core objective of increasing the accessibility of milled cereal grains with iron and folic acid consumed by the Brazilian population to reduce the prevalence of iron-deficiency and neural tube defects in Brazil (National Agency of Sanitary Surveillance, 2006).

However, iron-fortified wheat flour is not always available, or it is consumed in small quantities to be affective by poor children 6 to 60 months of age (Beinner & Lamounier, 2003). Fortification of specific foods, as part of a complementary diet, has shown to be more effective for the control and prevention of iron-deficiency among infants (Andrade, 2004). In addition, and according to Hurrell (1997) it is likely that the low levels of elemental iron added to wheat flour (40 mg/Kg) would have little impact on iron nutrition, but the much higher levels added to commercial infant cereals (200-550 mg/Kg) together with vitamin C, could contribute substantially to the prevention of iron deficiency anemia.

However, this measure becomes questionable in relation to infants, age of greatest risk for anemia due to the fact that these foods are not recommended and regularly consumed in sufficient quantities to meet the iron needs of this particular group. Moreover, it is likely that the low level of elemental iron (40mg/kg) added to wheat flour has little impact on nutritional status of children. No effect of flour fortification was observed in hemoglobin levels of children under five years in the city of Pelotas. Fact can partly be explained both by insufficient consumption of flour and also by the low bioavailability of dietary iron. The study was conducted between May and June 2004, prior to the mandatory fortification of flour and 12 and 24 months after the implementation of the action which occurred between 2005 and 2006 (Assunção et al., 2007). Moreover, in Brazil, there is not a monitoring program of mandatory fortification of flour.

4. Fortification of milk

The Brazilian Pediatrics Society has recommended the use of infant formula supplemented with iron for infants until the age of two as supplementary feeding with breastfeeding. However, cow's milk is an important food consumed by children especially those families of low socioeconomic status. Cow´s milk presents low bioavailability of iron, and consumption of excessive amounts of fresh or pasteurized cow's milk may be associated with occult intestinal blood loss during infancy, which may also contribute towards increasing the occurrence of anemia in infancy (Torres et al., 2000). The use of cow's milk, due to social-economic and cultural practices, is used frequently in Latin America, including Brazil, during infancy, and iron fortification of this vehicle is an inexpensive alternative to increasing iron levels in children (Torres et al., 1995).

Torres et al. (1995) studied the impact using powdered whole milk fortified with 9 mg of iron and 65 mg of vitamin C per 100 g during six months in 107 children in municipal daycare facilities, and another 228 children at public health clinics in the city of São Paulo. At baseline intervention, 66.4% and 72.8% of the children attending public daycare and public health clinics were diagnosed with anemia, respectively. At six months post study, the percentage of children still anemic decreased to 20.6% in daycare and 18% in children seen at health clinics. In a later study, Torres et al. (2000) evaluated the use of 3 mg of amino acid chelate in pasteurized cow's milk (3 mg/L). During the 12-month study, 239 children 6 to 42 months of age received, daily, one liter of fortified cow's milk. The mean hemoglobin levels at baseline for children less than 12 months, 12 to 23 months, 24 to 35 months, and 36 months of age, and older were 10.2 ± 1.3, 10.1 ± 1.6, 11 ± 1.3 and 11.8 ± 1.3 g/dl, respectively. At baseline, anemia prevalence was evaluated at 62.3%, and at six months, the percentage of children still anemic decreased to 41.8% and 26.4% after 1-year, respectively. Mean hemoglobin levels at 12 months were 11.1 ± 1.3, 11.6 ± 1.1, 12 ± 1.2, and 12.1 ± 1.0 g/dl, for 11, 12 to 23, 24 to 35, and 36 months of age, respectively. The increases were significant for the first three age groups, but not for the last group (36 months and older).

Braga (1996) evaluated 102 children aged two to six years of age from a low, socio-economic community, enrolled in municipally funded daycare facilities in the city of São Paulo. Using an infant formula, 14 mg of iron and 100 mg of ascorbic acid were added to 200 ml of formula daily during 180 days. At the conclusion of the study, significant increases were observed in anthropometric indices (not shown here), mean hemoglobin (Hb) levels and hematocrit (Htc) values at baseline (Hb: 12.1 ± 0.66 g/dl; Htc: 35.7 ± 1.9) and post study (Hb: 12.7 ± 0.66 g/dl; Htc: 37.9 ± 1.9) showed improvements. The authors concluded that the pre-school children could benefit in the control and prevention of anemia with a permanent iron-fortification program of foods in daycare facilities.

In another study to evaluate iron fortification of infant formula, Ferreira (2000) randomly assigned 111 children, between the ages of four and six months, to two intervention groups during six months: the experimental group (68 infants) received iron fortified (1.8 mg ferrous sulfate/200 ml) milk formula and a control group (43 infants) received milk formula (0.7 mg iron/200 ml). At baseline, anemia prevalence in groups 1 and 2 was 63.2% and 67.4%, respectively. Mean hemoglobin levels in group 1 increased from 10.6 g/dl to 11.3 g/dl, however, in group 2, mean Hb actually decreased from 10.6 g/dl to 10.1 g/dl at six months. Similar significant results were seen for mean ferritin values: at baseline, ferritin values increased from 34.8 to 44.8 mcg/dl, but in group 2, mean ferritin values decreased from 41.8 to 26.1 mcg/dl. Hb and ferritin status were significantly improved in iron fortified group. Overall, the anemia prevalence decreased from 63.2% to 33.8% in group 1, and increased from 67.4% to 72.1% in group 2.

It should be noted that effectiveness of iron fortified fresh or pasteurized cow's milk and milk formulas will depend on several factors such as iron compounds, quantity, bioavailability, iron enhancers and inhibitors likely to affect bioavailability, and overall added cost to the targeted consumer.

5. Iron fortification of biscuits and bread rolls

Some studies were conducted on the effect of bovine hemoglobin-fortified cookies on the hemoglobin levels of 16 iron-deficient preschool children in northeast Brazil (Nogueira et al., 1992). Each child was offered five cookies per day containing 1.25 mg of iron over three months as part of their normal school meal program. An evaluation of the total nutrients offered to the children showed an iron intake of just 4.0 mg/day. Baseline mean hemoglobin was 9.4 ± 2.6 g/dl, and after three months, mean hemoglobin increased to 13.2 ± 0.2 g/dl. Initial anemia prevalence was 73% and disappeared at three months post intervention. With the addition of bovine hemoglobin-fortified cookies to the children's diet, total iron intake increased to an average of 8.3 mg (83% of iron RDA – Recommended Daily Allowance) at a total cost of US$ 0.50 per child, with no measurable side effects or taste alterations reported.

A project developed with 1500 children from daycare centers in the city of Barueri, Sao Paulo, using cookies and breads fortified with iron aminoquelato at a dose of 2 mg / day, showed reduced levels of anemia from 32% to 11% in a period of 2 months of intervention, with positive change for the weight / height and height / age (Fisberg et al., 1996). Giorgini et al. (2001) evaluated 89 preschool children during six months in a study using iron bis-glycinate chelate. Children received two sweet rolls twice daily each fortified with 2 mg iron bis-glycinate (4 mg/day) five days a week. At baseline, 28% of the children had hemoglobin levels less than 11.0 g/dl, and at six months end study, nine percent of the children continued to be anemic. Mean hemoglobin at baseline was 11.5 g/dl, and at end, 12.6 g/dl. Mean hemoglobin increased 1.1 g/dl in non-anemic children and 1.4 g/dl in anemic ones. At the start of the study, mean ferritin level was 11.3 µg/l, and upon conclusion, mean ferritin increased significantly to 20.2 µg/l. Anthropometric indicators for weight/age and height/age also increased significantly.

However, the problem of fortification of breads and crackers is that these foods are not consumed in sufficient quantities to meet the needs of infants, and often not even part of the food habits of this age group at highest risk for anemia. Despite the universal assumption that biscuits and sweet rolls are consumed by almost everyone, biscuits and sweet rolls consumption by infants, toddlers and school children are quite different. As a consequence, the fractional iron intake contribution would be too low in a flour-based fortification program for infants. But these two vehicles – biscuits and sweet rolls complement each other, resulting in a significant reduction of the population below the iron RDAs (Vellozo et al., 2003).

6. Iron fortification of potable drinking water

The addition of iron to potable drinking water is one alternative to the control and prevention of iron deficiency and anemia. This rather simple method can reach a large part of the Brazilian population at each level of the social-economic stratum by the use of drinking water on a daily basis. Drinking water, other than used for drinking, is commonly used for preparation of foods, which may contribute even more towards increasing iron ingestion (Ferreira et al., 1991).

Dutra de Oliveira et al. (1994) evaluated 31 preschool children aged two to six years enrolled in daycare facilities in Ribeirao Preto, Sao Paulo. During eight months, children consumed iron-fortified drinking water (20 mg Fe/Liter) which resulted in a significant decrease in the prevalence of anemia. At baseline, anemia prevalence was diagnosed in 58% of subjects. At four months 16% continued anemic, but at eight months post-study intervention anemia virtually disappeared since anemia was present in only 3% of subjects. Mean hemoglobin levels at baseline (10.6 ± 1.1 g/dl) increased significantly to 12.1 ± 1.4 g/dl at four months, and 13 ± 1.1 g/dl at end study. In a later study, Dutra de Oliveira et al.(2002) studied low-income families during four months in which 21 families with children aged one to six years where divided into experimental and control groups. In the experimental group, family members consumed iron-fortified drinking water containing 10 mg of ferrous sulfate plus 60 mg of ascorbic acid per liter of water. The control group consumed their drinking water without the addition of iron or ascorbic acid. Results were very promising and showed that hemoglobin levels in children increased from 10.9 ± 1.1 g/dl to 11.7 ± 1.1 g/dl after four months of fortification intervention. Similar results were observed in the experimental adult group in which hemoglobin levels increased (12.9 ± 1.7 g/dl to 13.7 ± 1.7 g/dl). Results for ferritin were also positive in the experimental group in which ferritin levels increased in children, and significantly in adults. According to the authors, the iron fortification of drinking water is an effective, feasible alternative and practical way to distribute iron to low-income families, is technically inexpensive and has the promising potential for the control and prevention of anemia in Brazil and in other countries.

In another study, 160 preschool children from eight municipal daycare facilities benefited from daily consumption of iron (12 g element iron/L) plus ascorbic acid (90 mg/L) prepared in 20-L plastic water jugs (Beinner et al., 2005). Mean Hb at baseline and after eight months of intervention increased significantly from 11.8 ± 1.3 g/dl to 12.4 ± 0.93 g/dl, respectively. The prevalence of iron deficiency determined by hemoglobin levels decreased from 43.2% to 21% at eight months post intervention. Significant ($p< 0.05$) anthropometric growth indicators- weight/age, height/age and weight/height were also observed during the study. Fundamentally important to the success of this study was education of the targeted population, which resulted in behavior change and a greater awareness of the importance of combating iron deficiency and anemia by the use of iron-fortified drinking water.

The use of drinking water as a vehicle for the control and prevention of iron deficiency and anemia is an effective and efficient model, which can be used in targeting preschool children enrolled at daycare facilities, and/or at the household level, which will include all family members. Consumption of drinking water fortified with iron can contribute to increasing iron-intake to meet minimal Recommended Nutrient Intake (RNI) allowance of bioavailable iron acceptable to preschool children aged 6 to 59 months of age. Other then adding iron-concentrate to the appropriate number of liters of drinking water, it is easy to distribute and can be easily monitored.

In children attending daycare centers in Belo Horizonte city, southern of Brazil, a longitudinal study was conducted to evaluate the effectiveness of fortification of drinking water with iron and vitamin C in the reduction of the anemia as well as to identify the prevalence of anemia. It was evaluated 380 children aged six to 74 months. Since 55 did not participate in the second evaluation, a total of 312 children assessed before and at the end of the intervention. To study the identification of risk factors, it was evaluated only children under five years old, the group with the highest risk for anemia. A questionnaire was applied to parents or responsible for the children, containing information socioeconomic, maternal and related to the children's health. Anthropometric measurements (weight and height) and fingerstick blood samples occurred in two periods: before and after five months of fortification. Children were considered as anemic with hemoglobin < 11.0 g / dl for the group aged 6 to 59 months, and values < 11.5 g/dL for those aged 60 to 74 months. Multivariate analysis was performed to evaluate the association between these variables and anemia. The total number of children evaluated before and after the fortification was 318, being 52.2% male, with average of 45.4 ± 15.8 months. The prevalence of anemia decreased significantly from 29.3% before the fortification, to 7.9% at the end of the study ($p< 0,001$). Considering the prevalence by age group, a reduction of 62.5%, 75% and 78.8% was found for children of 24 months, 24 to 48 months and > 48 months, respectively. The hemoglobin median increased 10.2%: from 11.8 g/dL to 13 g/dL, with a significant increase in all age groups. There was improvement in height-for-age and weight-for-age, however, only the first measurement showed a significant difference. For the study of the risk factors of anemia, the prevalence of anemia in this population was 30.8%, and the prevalence was 71.1% in children aged ≤ 24 months. The risk factors of anemia were age ≤ 24 months (OR: 9.08 CI: 3.96 to 20.83), and height-for- age < -1 z score (OR: 2.1, CI: 1.20 to 3.62). The fortification of water with iron and vitamin C significantly reduced the prevalence of anemia in children attending daycare centers, as well as it improved the nutritional status of them, being considered an important strategy to control this nutritional deficiency (Rocha, 2010)

7. Iron fortification of bean and rice

In southeastern state of São Paulo, Brazil, studies were carried out, during four months, to evaluate bean flour enriched with iron in 85 anemic children two to five years of age. Results demonstrated a non-significant increase in anthropometric measurements and a significant reduction in the prevalence of anemia, which at baseline, was 13%, and at end study, anemia had disappeared in subjects that had received the iron-fortified bean flour (Fisberg et al., 2003). Unfortunately, milled bean flour represents a greater cost burden, and in addition, is not widely consumed throughout Brazil.

Rice is another alternative for food fortification. One study was conducted in four nurseries in Rio de Janeiro (RJ), with children in the intervention group (n = 180) attending two nurseries and the control group (n = 174) in the other two nurseries. It was observed an increase in hemoglobin concentration in both groups, with the reduction in the prevalence of anemia in the intervention group was 37.8% to 23.3% and for the control group was 45.4% for 33, 3%, with no difference in reduction between the groups. According to the authors, the total amount of iron available was not sufficient to achieve more significant results in the intervention group, after four months of study (Bagni et al., 2009). The other study was conducted with families in the metropolitan area of Belo Horizonte. A group of 84 children received iron-fortified rice (23 mg Fe / day) and another group received ferrous sulfate (25 g Fe / L). After five months of intervention, there was a reduction in the prevalence of anemia in both groups, with an initial prevalence of 100% in both groups, decreasing to 61.9% for the group receiving the fortified rice and 85.6% for the group receiving ferrous sulfate, with a significant difference between groups (Beinner et al., 2009). Regarding rice, more studies are needed to evaluate the timing and dose required fortification of that vehicle, to achieve preventive effects and / or significant curative as well as assess the effect of simultaneous use with other supplements containing iron

8. Other iron fortified foods

Orange juice fortification studies shown improvement in childhood anemia. In ongoing studies with iron fortification of foods, De Paula & Fisberg (2001) evaluated the use of 20 g of iron fortified sugar added to orange juice offered to 93 preschool children during six months. Children were divided into two groups: group 1 received 10 mg of iron per kilo of sugar, and group 2 received 100 mg of iron per kilo of sugar, both in the form of ferrous trisglycinate. Anemia prevalence in both groups evaluated at baseline was 38.1% and 29.4%, respectively. At six months post study intervention, anemia prevalence in both groups decreased to 19.7% and 19.6%, respectively ($p = 0.01$). Mean hemoglobin levels increased to 0.4 g/dl; in anemic children alone, mean Hb increased greatly to 1.3 g/dl and 1.5 g/dl in groups 1 and 2, respectively ($p < 0.001$). According to ferritin results, there was a positive trend towards normalization of ferritin values in iron-deficient children. It was suggested, in terms of cost, that use of 10 mg iron/Kg be used when compared with 100 mg/Kg, as same results were observed.

In yet another study using orange juice as an iron vehicle, Almeida et al. (2003) evaluated iron fortification of this widely produced fruit rich in vitamin C, which greatly facilitates iron absorption. Fifty preschool children consumed orange juice with iron (10 mg ferrous sulfate per 100 ml of concentrated orange juice) twice daily during four months. Anemia prevalence decreased from 60% to 20% at end study, and mean hemoglobin level increased from 10.5 ± 1.7 to 11.6 ± 1.1 g/dl ($p = 0.00$). The use of iron fortified orange juice is a promising strategy as a complimentary vehicle for ingestion of iron in children. Orange juice is widely consumed by all levels of the social strata in Brazil. An iron compound can be added during processing without provoking organoleptic changes (i.e., color, flavor, and consistency), and even allow for much higher quantities of iron- from 3 to 10 times more- than in other targeted or mandatory foods. The added cost can be absorbed through advertising and processing.

The manioca flour enriched with ferrous bis-glycinate was studied during four months in 80 preschool children enrolled in a philanthropic institution in the city of Manaus. Anemia prevalence decreased significantly from 22.7% to 8% after four months of intervention ($p < 0.05$). According to the authors, mandioca flour is widely consumed in the North region of the country and can be considered a promising food vehicle in the control and prevention of iron deficiency and anemia (Tuma et al., 2003)

Studies on fortification showed a positive response, both in relation to acceptance of fortified food, and prevention, as in the recovery of hemoglobin levels in both groups (Hertrampf, 1990; Torres et al., 1995; Vitolo et al., 1998). The food industries have used the enrichment of their products as a commercial appeal, focused on creating a quality attribute to further enhance the marketing of their products. However, there is no data in Brazil to assess the impact of these foods, fortified voluntarily by industry, in the prevalence of anemia.

9. Conclusion

Studies on iron fortification of foods, over the last twenty years, have shown promising results in the control and prevention of iron deficiency and anemia in infant and child populations. Unfortunately, only a small number of efficacy and effectiveness trails of iron fortification of foods and liquids conducted in Brazil have been published. Researchers have used various types of food vehicles as well as different iron compounds in attempt to reduce nutritional deficiency, particularly an iron deficiency.

The high prevalence of iron deficiency and anemia in infancy in most regions of Brazil have called attention to an inadequate nutrition making this a serious public health problem leading to eventual losses in terms of future growth and productivity at all stages of human development. State and federal governmental health agencies must move forward to prioritize national nutrition agenda that will draft mandatory fortification of food staples for mass consumption. Finally, fortified food is made available to vulnerable populations when industry is motivated to develop the logistics needed to fortify their products and when government is motivated to change policy requiring fortification.

10. References

Agência Nacional de Vigilância Sanitária. Ferro e ácido fólico enriquecem farinha. Available from: http://www.anvisa.gov.br/divulga/noticias/2004/180604.htm. Accessed in 2006 (Dec 18).

Andrade, KC. (2004) A fortificação de alimentos com ferro no controle da anemia ferropriva. *Revista Brasileira de Ciências da Saúde*, Vol.II, No.3, pp. 50-5.

Assao TY, Silva DG, Ribeiro LC, Devincenzi MU, Sigulem DM. (2004) A importância do ferro na saúde e nutrição do grupo materno-infantil. *Compacta Nutrição*, Vol.5, No.3, pp. 7-22.

Assunção MC, Santos IS, Barros AJ, Gigante DP, Victora CG. (2007) Efeito da fortificação de farinhas com ferro sobre anemia em pré-escolares, Pelotas, RS. *Rev Saúde Pública*, Vol.41, No.4, pp. 539-48.

Bagni UV, Baião MR, Santos MMAS, Luiz RR, Veiga GV. (2009) Efeito da fortificação semanal do arroz com ferro quelato sobre a frequência de anemia e concentração de hemoglobina em crianças de creches municipais do Rio de Janeiro, Brasil. *Cad. Saúde Pública*, Vol.25, No.2, pp. 291-302.

Beinner MA, Lamounier JA. (2003) Recent experience with fortification of foods and beverages with iron for the control of iron-deficiency anemia in Brazilian children. Food Nutr Bull. Vol.24, No.3, pp. 268-74.

Beinner MA, Lamounier JA, Tomaz C. (2005) Effect of iron-fortified drinking water of daycare facilities on the hemoglobin status of young children. *Journal of the American College of Nutrition*, Vol.24, No.2, pp. 107-114.

Beinner MA,Velasquez-Melendez G, Pessoa MC, Greiner T. (2009) Iron-fortified Rice is as efficacious as supplemental iron drops in infants and Young children. *Journal of Nutrition*, Vol.4, pp. 1-5.

Braga JAP. Avaliação do estado nutricional de crianças submetidas a intervenção nutricional com um suplemento lácteo fortificado com ferro. [These] São Paulo (SP): Universidade Federal de São Paulo; 1996

Capanema FC, Lamounier JA, Rocha DS, Tonidandel WC, Drumond CA, Maia LS, Gomes CMA. (2008) Prevalência de anemia em crianças pertencentes a creches de Belo Horizonte – MG. *Revista da Fundação Ezequiel Dias*, Vol.2, No.1, pp. 75-82.

Cardoso MA, Penteado MVC. (1994) Intervenções nutricionais na anemia ferropriva [Nutritional strategies for controlling iron deficiency anemia]. *Cad Saúde Pública = Rep Public Health.*,Vol.10, No.2, pp. 231-40.

Castro TG et al. (2011) Anemia e deficiência de ferro em pré-escolares da Amazônia Ocidental brasileira: prevalência e fatores associados. *Cad. Saúde Pública*, Vol. 27, No. 1, pp. 131-42.

de Paula RA, Fisberg M. (2001) The use of sugar fortified with iron tris- glycinate chelate in the prevention of iron deficiency anemia in preschool children. *Arch Latinoam Nutr.;* Vol.51, Suppl 1, pp. 54-9.

de Almeida CAN, Crott GC, Ricco RG, Del Ciampo LA, Dutra-de-Oliveira JE, Cantolini A. (2003) Control of iron-deficiency anaemia in Brazilian preschool children using iron-fortified orange juice. *Nutrition Research*, Vol. 23, No. 1, pp. 27-33.

Dutra-de-Oliveira JE, Ferreira JB, Vasconcellos VP, Marchini JS. (1994) Drinking water as an iron carrier to control anemia in preschool children in a day-care center. *J Am Coll Nutr*, Vol. 13, No. 2, pp. 198-202.

Dutra-de-Oliveira JE, de Almeida CA. (2002) Domestic drinking water--an effective way to prevent anemia among low socioeconomic families in Brazil. *Food Nutr Bull.*, Vol. 23, (3 Suppl), pp. 213-6.

Ferreira AMA. Prevenção da anemia ferropriva em lactentes que freqüentam creches do município de São Paulo, através de uma fórmula láctea infantil fortificada com ferro.[These] São Paulo (SP): Universidade Federal de São Paulo; 2000.

Ferreira JF, Aranda RA, Bianchi MLP, Desai ID, Dutra-de-Oliveira JE (1991) Utilização da água potável como veiculo de nutrientes: estudos experimentais com ferro [Use of drinking water as a vehicle of nutrients: experimental studies with iron]. *Arch Latinoam Nutr.*, Vol. 41, No. 3, pp. 400-8.

Fisberg M, Velloso EP, Ribeiro RMS, Zullo M, Braga JAP, Soraggi C, Kliamca PE, Chedid EA, Schuch M, Valle J, Cardoso R, Krumfli M, Graziani E. (1996) - Projeto Barueri. *Pediatr. Atual*, Vol. 11, No. 4, pp. 19-26.

Fisberg M, Lima AM, Naufel C, Rodrigues C, Rhein SO, Oliveira S. (2003) Feijão enriquecido com ferro na prevenção de anemia em pré-escolares. *Nutrição em Pauta*, Vol. 59, No. 11, pp. 10-18.

Fontes F. (2007) Politicas de combate de carencias de micronutrientes em Panamá. *Rev Méd de Minas Gerais*, Vol. 17, pp. S74-9.

Giorgini E, Fisberg M, De Paula RAC, Ferreira AMA, Valle J, Braga JAP. (2001) The use of sweet rolls fortified with iron bis-glycinate chelate in the prevention of iron deficiency anemia in preschool children. *ALAN*, Vol.51, 1 Suppl, pp. 48-52.

Herrera MP. (2007) Estrategias y acciones para combatir la anemia y la deficiencia de hierro: la experiência de Cuba en la fortificación de alimentos con hierro. *Rev Méd Minas Gerais*, Vol.17, pp. S86-9.

Hertrampf E, Olivares M, Walter T, Pizarro F. (1990) Anemia ferropriva en el lactante: erradicación con leche fortificada con hierro. *Rev Méd Chile.*; Vol.118, pp. 1330-7.

Hurrell RF. (1997) Preventing iron deficiency through food fortification. *Nutr Rev.*, Vol.55, No. 6, pp. 210-22.

Instituto Brasileiro de Geografia e Estatística – IBGE Pesquisa de Orçamentos Familiares 2008-2009 Antropometria e Estado Nutricional de Crianças, Adolescentes e Adultos no Brasil, 2010.

Jordão RE, Bernardi JLD, Filho AAB. (2009) Prevalence of iron-deficiency anemia in Brazil: a systematic review. *Rev Paul Pediatr*, Vol.27, No. 1, pp. 90-8.

Looker AC, Dallman PR, Carroll MD, Gunter EW, Johnson CL. (1997) Prevalence of iron deficiency in the United States. Prevalence of iron deficiency in the United States. *J Am Med Assoc.*, Vol.277, No. 12, pp. 973-6.

Matta IEA, Veiga GV, Baião MR, Santos MMAS, Luiz RR. (2005) Anemia em crianças menores de cinco anos que freqüentam creches públicas do município do Rio de Janeiro, Brasil. *Rev. Bras. Saúde Matern. Infant.*, Vol.5, pp. 349-57.

Morais MB, Alves GM, Fagundes-Neto U. (2005) Nutritional status of Terena indian children from Mato Grosso do Sul, Brazil: follow up of weight and height and current prevalence of anemia. *J Pediatr*, Vol.81, pp. 383-9.

Neves MBP, Silva EMK, Morais MB. (2005) Prevalence and factors associated with iron deficiency in infants treated at a primary care center in Belém, Pará, Brazil. *Cad. Saúde Pública*, Vol.21, No. 6, pp. 1911-18.

Nilson A, Piza J. (1998) Food fortification: a tool for fighting hidden hunger. *Food and Nutrition Bulletin*, Vol.19, No. 1, pp. 49-60.

Nogueira NN, Colli C, Cozzolino SMF. (1992) Controle da anemia ferropriva em pré-escolares por meio da fortificação de alimento com concentrado de hemoglobina bovina: estudo preliminar [Iron deficiency anemia control in pre-school children by food fortification with bovine hemoglobin: preliminary study]. *Cad Saúde Pública = Rep Public Health.* Vol.8, No. 4, pp. 459-65.

Rocha DS. Efetividade da fortificação da água com ferro e vitamina c na prevalência de anemia e no estado nutricional em crianças assistidas em creches De Belo Horizonte-MG, 2010. Doutorado [These] - Universidade Federal de Minas Gerais

Sharieff W, Horton S, Zlotkin S. (2006) Economic Gains of a home fortification program: evaluation of "Sprinkles" from provider's perspective. *Canadian Journal of Public Health*, Vol.97, No. 1, pp. 20-23.

Sherry B, Mei Z, Yip R. (2001) Continuation of the decline in prevalence of anemia in low-income infants and childrens in five states. *Pediatrics*, Vol.107, No. 4, pp. 677-82.

Sprinkles Global Health initiative: about sprinkles.[cited 2009 May]. Available from: <http://www. sghi.org/index.html>.

The Global Burden of Disease: 2004 updated. Part 3: Disease incidence, prevalence and disability. World Health Organization Press, Switzerland, 2008. Pags 31-32. Available from: http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf

Torres MAA, Souza Queiroz S. (2000) Prevenção da anemia ferropriva em nível populacional: uma revisão da literatura dos últimos quinze anos [Prevention of iron deficiency anemia in public health: a fifteen years review of the literature]. *Nutrire*. Vol.19, No. 20, pp. 145-64.

Torres MAA, Sato K, Lobo NF, Queiroz SS. (1995) Efeito do uso de leite fortificado com ferro e vitamina C sobre os níveis de hemoglobina e condição nutricional de crianças menores de 2 anos [Effects of vitamin C and iron-fortified milk use on hemoglobin levels and nutritional condition of children cared for in day]. Rev Saúde Pública = J Public Health, Vol.29, No.4, pp. 301-7.

Tuma RB, Yuyama LKO, Aguiar JPL, Marques HO. (2003) Impacto da farinha de mandioca fortificada com ferro aminoácido quelato no nível de hemoglobina de pré-escolares [Impact of cassava flour fortified with iron amino acid chelate on the hemoglobin level in pre-schools]. *Rev Nutr,* Vol.16, No.1, pp. 29-39.

Vellozo EP, Silva R, Fagioli D. (2003) Pão enriquecido com ferro na prevenção da anemia de crianças matriculadas em creches da Prefeitura do Município de São Paulo. *Nutrição em Pauta*. Vol.63, pp. 32-42.

Vitolo MR, Aguirre NC, Kondo MR, Giuliano Y, Ferreira N, Lopez FA. (1998) Impacto do uso de cereal adicionado de ferro sobre os níveis de hemoglobina e a antropometria de pré-escolares. *Rev Nutr.*; Vol.11, No.2, pp. 163-71.

World Health Organization. Preventing and Controlling Iron Deficiency Anaemia through Primary Health Care. A guide for health administrators and programme managers. Geneve: WHO; 1989.

Economic Stressors and Childhood Obesity: Differences by Child Age and Gender

Steven Garasky[1], Craig Gundersen[2],
Susan D. Stewart[3], Joey C. Eisenmann[4] and Brenda J. Lohman[3]
[1]*IMPAQ International, LLC,*
[2]*University of Illinois,*
[3]*Iowa State University,*
[4]*Michigan State University*
USA

1. Introduction

Childhood obesity is a public health challenge in the United States (U.S.) and elsewhere in the world. Additionally, those who are obese are heavier than in the past (Anderson & Butcher, 2006). In the U.S., one in three children is overweight or obese (Ogden et al., 2010), a prevalence that has tripled since 1970 (Anderson & Butcher, 2006; Kumanayika & Grier, 2006; Wang & Zhang, 2006). In response to this public health issue, Healthy People 2010 (US/DHHS, 2000) and President Obama (US/Office of the President, 2010) have identified childhood obesity as a national health priority as it has immediate consequences for a child's physical and psychological health (Puhl & Latner, 2007; Raman, 2002; Strauss, 1999; US/DHHS, 2000), as well as implications for future health (Freedman et al., 2007; Raman, 2002; Strauss, 1999; US/DHHS, 2000). Beyond negative health outcomes, there are also economic costs (e.g., greater need for health care) associated with childhood obesity (Marder & Chang 2006; Skinner et al., 2008). Thus, identifying factors related to childhood obesity not only has implications for the health and quality of life of children, but it also has important implications for family expenditures and health care costs.

It is commonplace to focus on physical inactivity and dietary factors as the cornerstones of the childhood obesity epidemic, but stress is another common feature of the landscape facing American families today. Stress manifests itself across numerous dimensions at both the individual and family level. While stress can be managed successfully by many individuals and families, in some cases stress can become severe enough to lead to serious health consequences. A vast literature has demonstrated the effects of stress on numerous health outcomes for children and adults (e.g., Dearing et al., 2006; Evans & English, 2002; Gee & Walsemann, 2009; Kort-Butler, 2009; Schilling et al., 2008), including childhood obesity (Eisenmann, 2006; Garasky et al., 2009; Gundersen et al., 2008; Lohman et al., 2009).

While this work has provided policymakers with important insights, a central issue related to this research must be addressed. Through regression analyses this work has found positive associations between stress and childhood obesity after controlling for a host of potential confounding factors, but it has not ascertained whether unobserved factors correlated with stress (e.g., a child's ability to cope with adverse conditions) may be the cause of the association. That is, these studies have implicitly assumed that time invariant unobserved factors were similar between different children and have ignored the possibility that children in households experiencing stress are different in unobserved ways from children in households not experiencing stress. To address this issue, the current study examined panel data using fixed effects models that controlled for time-invariant differences between children by using only within-individual variation to estimate the regression coefficients (Allison, 2005).

We examined data from the first two waves of the Child Development Supplement (CDS) of the Panel Study of Income Dynamics (PSID). Our results indicated that exposure to housing-related economic stressors leads to a higher probability of a child being obese and to higher levels of obesity, especially for younger females. Other forms of stress examined here (financial and neighborhood) were not related to child weight status. These findings were robust across a range of model specifications and suggest that efforts to reduce housing stress may also lead to reductions in childhood obesity.

2. Background

2.1 Theoretical framework

Child health is an important aspect of family well-being. The ecological theory of human development identifies four levels of influence faced by families attempting to maximize their well-being (Bronfenbrenner & Morris, 1998). These levels are the microsystem (e.g., individuals and families), the mesosystem (e.g., neighborhoods and social networks), the exosystem (e.g., community), and the macrosystem (i.e., larger cultural context). Theoretical work on the "stress process" defines *stress* as a negative physiological response and *stressors* as the external factors that cause this negative response (e.g., Aneshensel, 1992; Boss, 1988; Chrousos & Gold, 1992; Pearlin, Menaghan, Lieberman & Mullan, 1981; Pearlin, Schieman, Fazio & Meersman, 2005).

While genetic factors have consistently been shown to be central to whether a child is obese (e.g., Crossman et al., 2006; Gibson et al., 2007), environmental factors at the microsystem and mesosystem levels such as family, parental, and economic influences matter as well (e.g., Anderson et al., 2003; Cutler et al., 2003; Dietz & Robinson, 2008; Loureiro & Nayga, 2005). Environmental factors combined with a biological predisposition toward obesity provide the conditions for one's propensity for obesity to come to fruition (Anderson et al., 2003).

2.2 Stress and health

Stress is an environmental factor that often leads to reduced psychological and physiological health. One of the most common consequences of stress exposure is psychological distress, especially depression (e.g., Brooks et al., 2002; Kort-Butler, 2009; Schilling et al., 2008). Daily hassles and chronics stressors are the typical antecedents to depressive symptoms with stress leading to anxiety and aggression (Evans & English, 2002; Krause et al., 2003; Kort-Butler, 2009).

Physiologically, higher levels of stress have been associated with higher levels of self-reported illness (Gee & Walsemann, 2009; Goodman, 1999; Larson et al., 2008). Early and chronic exposure to stressors has been linked with cardiovascular disease and obesity.

Specifically, the psychosocial stressors that provoke exaggerated cardiovascular reactivity may also trigger overeating (Marniemi et al., 2002; Siervo et al., 2009). Indeed, research has found that stress exposure leads to increased cortisol levels, which enhances appetite and decreases leptin resistance, which increases the risk of obesity (Bjorntorp, 2001; Koch et al., 2008; Siervo et al., 2009).

The effects of economic stressors on health and well-being have been examined under a range of headings including economic stress, financial stress, economic hardship, economic strain, economic pressure, and material hardship (e.g., Fletcher et al., 2005; Kim & Garman, 2003; Kim et al., 2006). Lower income children are at an elevated risk of exposure to varied forms of these stressors (Gershoff et al., 2007; Pearlin et al., 2005; Wickrama et al., 2007) and may have more severe health consequences stemming from stress due to their greater exposure and vulnerability (Sampson et al., 1997; Spencer, 2001). Recently, scholars have shown an association between household- and individual-level indicators of stress and childhood obesity (Garasky et al., 2009; Gundersen et al., 2008; Lohman et al., 2009).

2.3 Stress and childhood obesity

Recent evidence suggests that stress faced by family members may lead to childhood obesity. Gundersen et al. (2008) found using data from the National Health and Nutrition Examination Survey (NHANES) that higher values of an index measuring cumulative stress exposure led to higher probabilities of obesity for food secure children in comparison to food insecure children. This result held for younger children, but the effect was statistically insignificant for older children. Lohman et al. (2009) found using data from the Three-City Study that higher levels of individual stress experienced by a child between the ages of 10 and 15 were statistically significantly associated with higher probabilities of childhood obesity. In contrast to Gundersen et al. (2008), they found that food insecure children with higher levels of maternal stress had higher probabilities of childhood obesity in comparison to food secure children. In should be noted that Gundersen et al. (2008) and Lohman et al. (2009) examined data from different surveys and employed different protocols for measuring food insecurity and stress which may have led to their finding somewhat conflicting results. Garasky et al. (2009) found using data from the PSID that stress was associated with childhood obesity, but the type of stress had differential effects for younger and older children. For younger children, lack of cognitive stimulation and emotional support in the household were associated with higher probabilities of obesity. However, for older children, mental and physical health problems and financial strain in the household were associated with higher probabilities of obesity. Finally, van Jaarsveld et al. (2009) conjectured that perceived stress in pre-adolescence may set adiposity trajectories, with no accentuation of differences due to perceived stress in adolescence, when they did not find an association between perceived stress and weight gain among British adolescents.

Important to this research, few studies have investigated whether associations between stressor exposure and weight status differ between females and males, nor have they examined whether the associations differ by gender within age groups (an exception being van Jaarsveld et al. (2009) who examined potential differences among adolescents by gender). Among adolescents, boys and girls have the same likelihood of being overweight (Anderson et al., 2003) with no differences in the effect of perceived stress on weight gain (van Jaarsveld et al., 2009). However, some research indicates that girls and boys cope with stress differently (e.g., Frydenberg & Lewis, 2004; Rudolph, 2002), although others have found no gender differences in adolescent coping (Wadsworth & Compas, 2002). Different coping styles for boys and girls may affect their risks of obesity. For example, eating-related responses to stress differ for boys and girls (Mikolajczyk et al., 2009) with adolescent girls being more likely than boys to suffer from disordered eating (Hepworth, 2004). Additionally, greater consumption of sweets and fast food consistent with higher rates of "emotional eating" has been found among women relative to men (Larsen et al., 2006).

The current study expanded on previous work by considering three types of economic stressors most commonly examined in the literature and most relevant to the home environments of children – housing, financial and neighborhood stressors – within a fixed effects framework. We examined separate groups of younger and older children and, given the mixed results by age and gender discussed above, compared the relationship between stress and obesity for girls versus boys within each age group. Previous studies did not systematically examine the potential effects of these economic factors on a child's propensity to be obese in this way. Additionally, previous work concentrated on binary measures of obesity (an exception being van Jaarsveld et al. (2009)). We utilized the obesity gap (Garasky et al., 2009; Jolliffe, 2004) to depict the extent of a child's obesity. A central advantage to using the gap measure is that it addresses a key disadvantage associated with binary measures of child weight status. A binary measure of obesity treats all children with a body mass index (BMI) ≥ 95th percentile for age and sex the same. With the obesity gap, these children are treated differently within the context of the models. From a policy perspective, these additional analyses are important as one may be especially interested in children with relatively high levels of obesity (US/Office of the President, 2010).

3. Methods

We controlled for time invariant unobserved factors through the following fixed effects model:

$$\mathrm{OB}_{it}^{\alpha} = \boldsymbol{\gamma}\mathbf{I}_{it} + \boldsymbol{\lambda}\mathbf{Y}_{it} + \mu_i + \varepsilon_{it} \tag{1}$$

where i denotes a child, t denotes the interview wave (t=1 or 2), $\mathrm{OB}_{it}^{\alpha}$ denotes the measures of the weight status of the child (in manners described below), **I** is a vector of the three economic stressor indices described below, **Y** is a vector of time varying covariates, μ is a child-specific fixed effect, and ε is an error term. We estimated logit fixed effects models for the binary specification of weight status and linear regression fixed effects models for our continuous (i.e., gap) measure. With respect to the direction of the influence of the economic stressors on childhood obesity, these models implicitly assumed that the stressors affected weight status. While in theory the relationship between the economic stressors and weight status could be bidirectional, our model was consistent with the vast majority of research in this area and every study discussed above.

3.1 Data

Our analyses were conducted with data from the first two waves of the Child Development Supplement (CDS-I and CDS-II) of the PSID. The PSID, begun in 1968, is a longitudinal study of a nationally representative sample of U.S. individuals and the families in which they reside. In 1997, a refresher sample of post-1968 immigrant families and their adult children was introduced to keep the study representative of the U.S. population (PSID, 2005). Currently, PSID interview waves are conducted biannually.

The CDS, a research component of the PSID focusing on children age 0-12 years in PSID families was introduced in 1997. The CDS examines a range of developmental outcomes within the context of family, neighborhood, and school environments. Assessments of cognition, behavior, and health status are obtained from a variety of sources including the child and the child's primary caregiver. The multi-method CDS survey design includes computer-assisted personal interviews (CAPI) and audio computer-assisted self-interviews (ACASI) (PSID, 2008b). Trained personnel measure the child's height without shoes using a rafter's square and tape measure. The child's weight is measured using a digital scale (PSID, 2008a). In 2002-2003, interviewers recontacted families that participated in CDS-I and remained active in the PSID as of 2001. Of those families, 91% were successfully reinterviewed (PSID, 2008b).

Information from CDS-I and CDS-II constituted the basis of our study. We supplemented these data with income and household composition data from contemporaneous PSID interview waves. Together, these data were well-suited for this analysis as they provided a large sample of households and detailed longitudinal information on child characteristics, family stressors, and relevant covariates. This data set had other strengths as well. In particular, the use of directly measured child height and weight strengthened this study as other large, national studies (e.g., Youth Risk Behavior Survey; National Longitudinal Survey of Youth) tend to rely on less reliable self-report or parental-report methods.

Our analytic sample consisted of 1,263 youths who at the time of the CDS-I interview were between 2 and 14 years of age. This research examined the full analytic sample, as well as two age-based subsamples. Consistent with other research, the full sample was split at age 8 years at the time of the CDS-I interview. Children less than 8 years of age (i.e., 24-95 months of age) at the time of the CDS-I interview were the younger sample (n = 677 children). Children at least 8 years of age (i.e., 90-167 months of age) at the time of the CDS-I interview comprised the older sample (n = 586 children). Youths under age 2 years (less than 24 months) at the time of the CDS-I interview and over age 18 years (over 228 months) at the time of the CDS-II interview were removed. Also, youths classified as underweight (body mass index < 5th percentile for age and gender as defined below) at either their CDS-I or CDS-II interview were removed from the sample since underweight children were not the focus of the referent group (i.e., normal weight youth).

3.2 Variables and descriptive statistics

3.2.1 Dependent variables

The measures used to delineate child weight status began with the calculation of a child's body mass index (BMI, kg/m^2). The BMI for each child was then mapped into a percentile based on age (in months) and gender using the Centers for Disease Control and Prevention (CDC) growth charts for the United States (e.g., Kuczmarski et al., 2002; Ogden et al., 2002). Our measures of obesity (OB^{BIN} and OB^{GAP}) were derived from these BMI percentiles and employed the definitions of the American Academy of Pediatrics (Barlow, 2007). That is, we set the obesity cutoff at the 95th percentile for age and gender.

Our binary measure of child weight status (OB^{BIN}) was defined as follows:

$$\begin{aligned} OB^{BIN} &= 1 \text{ if } BMIPER \geq 95 \\ OB^{BIN} &= 0 \text{ otherwise} \end{aligned} \qquad (2)$$

where BMIPER was the child's BMI percentile for age and gender.

Our measure of obesity severity (OB^{GAP}, the obesity gap measure) was defined as follows:

$$OB^{GAP} = [(BMIPER - 95) / 5] \text{ if } BMIPER \geq 95 \quad (3)$$
$$OB^{GAP} = 0 \text{ otherwise}$$

To better understand how OB^{BIN} and OB^{GAP} were defined, consider three children who have BMI percentiles for age and gender of 50, 96 and 99, respectively. Regarding OB^{BIN}, the first child, and any child with a BMIPER below the 95th percentile, will have OB^{BIN} set equal to 0. For the second and third child, and any child with BMIPER ≥ 95th percentile, OB^{BIN} will equal 1. As for OB^{GAP}, the first child, and any child with a BMIPER below the 95th percentile, OB^{GAP} will be set to 0 as well. Key to these analyses is that the second and third children, both of whom are classified as obese, will have different values for OB^{GAP}. OB^{GAP} will equal 0.2 for the second child, while OB^{GAP} will equal 0.8 for the third child. That is, children with more severe obesity (higher BMI percentiles for age and gender) will have higher obesity gap scores.

Variables	All Children 2 to 14 yrs		Younger Children 2 to <8 yrs		Older Children 8 to 14 yrs	
	CDS-I	CDS-II	CDS-I	CDS-II	CDS-I	CDS-II
Dependent Variables						
Obese	0.200	0.210	0.241	0.230	0.154	0.186
Obesity gap	0.130 (0.291)	0.120 (0.268)	0.170 (0.330)	0.136 (0.283)	0.084 (0.228)	0.101 (0.249)
Stressor Indices						
Housing stressors (range = 0-4)	0.245 (0.514)	0.196 (0.484)	0.270 (0.554)	0.225 (0.507)	0.217 (0.463)	0.164 (0.453)
Financial stressors (range = 0-10)	1.451 (1.529)	1.182 (1.496)	1.539 (1.530)	1.205 (1.513)	1.348 (1.522)	1.155 (1.479)
Neighborhood stressors (range = 0-2)	0.299 (0.584)	0.295 (0.592)	0.298 (0.575)	0.301 (0.609)	0.300 (0.595)	0.289 (0.573)
Control Variables						
Number of persons in household	4.154 (1.200)	4.064 (1.217)	4.049 (1.216)	4.165 (1.133)	4.275 (1.171)	3.947 (1.299)
Annual total family income ($1000)	50.565 (54.930)	70.725 (106.238)	47.144 (44.570)	65.316 (71.196)	54.516 (64.687)	76.974 (135.712)
Number of respondents	1263		677		586	

[a] Percent of sample reported for categorical measures. Means with standard errors reported in parentheses for continuous measures. Children are classified as 'younger' or 'older' based on their age at the time of the CDS-I interview. Obese (BMI ≥ 95th percentile).

Table 1. Summary statistics: Variables in multivariate regressions.[a]

Summary statistics for the variables in the multiple regressions for each wave are reported in Table 1 with changes in the variables across interviews reported in Table 2. About one-fifth of the analytic sample was categorized as obese in each wave. The group of younger children had a higher proportion that was obese compared to the older children. About one-fourth (22.7%) of all of the children experienced a change in weight status between interviews with a higher percentage of younger children (28.5%) experiencing change compared to the older children (15.9%). Older children, however, were more likely to become obese (9.6%) between interviews than to no longer be obese (6.3%) at the second interview. Younger children slightly more frequently became non-obese between waves (13.7% became obese, 14.8% were no longer obese at CDS-II). By definition, there was more change in weight status when status was assessed via the obesity gap (31.8% for all children, 37.8% for younger children, and 24.9% for older children) than when obesity was assessed through the standard binary measure.

Variables	All Children 2 to 14 yrs		Younger Children 2 to <8 yrs		Older Children 8 to 14 yrs	
	Increase	Decrease	Increase	Decrease	Increase	Decrease
Dependent Variables						
Obese[a]	11.8	10.9	13.7	14.8	9.6	6.3
Obesity gap	17.0	14.8	18.0	19.8	15.9	9.0
Stressor Indices						
Housing stressors	10.4	15.5	11.7	15.8	8.9	15.0
Financial stressors	25.7	36.9	25.4	39.3	26.1	34.1
Neighborhood stressors	13.3	13.1	12.9	12.9	13.8	13.5
Control Variables						
Number of persons in household	22.3	28.9	29.8	21.0	13.7	38.1
Annual total family income ($1000)	78.4	21.1	77.7	21.9	79.2	20.3
Number of respondents	1263		677		586	

Children are classified by their age at the time of the CDS-I interview. Obese (BMI ≥ 95th percentile).

[a] Increase defined as becoming obese between CDS-I and CDS-II interviews. Decrease defined as no longer being obese at CDS-II interview.

Table 2. Summary statistics: Change in variables in multivariate regressions from CDS-I to CDS-II (% of children).

3.2.2 Economic stressors

Stress is a difficult concept to measure. Both subjective (i.e., directly asking individuals how much stress they feel from a given stressor) and objective (i.e., observing stressors within an individual's environment or assessing physiologic markers such as cortisol) assessments of stress have been used in research. For example, Gottholmseder et al. (2009) examined the effect of commuting on perceived stress via a survey question asking "How do you feel when you arrive at your place of work (under normal traffic conditions)?" van Jaarsveld et al. (2009) when examining the link between stress and weight gain among adolescents assessed perceived stress with four questions that focused on stress and coping over the preceding month (e.g., "How often have you felt that you could not control the important things in your life?"). On the other hand, Kim et al. (2006) linked the stress caused by excessive debt (an objective financial stressor) to a higher likelihood of workplace absenteeism.

This study employed objective measures of stress for three categories of economic stressors (housing, financial and neighborhood). Consistent with individuals experiencing "stress pile-up" as a result of dealing with multiple stressors at once (McGuigan, 1999; White & Klein, 2002) and aggregate economic risk being a more important correlate than any single economic risk (MacFadyen et al., 1996), a stressor index was calculated for each category by summing the dichotomous response values for the variables in the category.

Housing stressors. Economic stressors related to housing were measured via four variables. CDS respondents when asked about economic problems in the last 12 months were asked three questions about housing experiences. "Did they move to cheaper living quarters?" "Did they move in with other people?" "Did they send one or more of the children to live with someone else?" A 'yes' response to any of these items was considered an indicator of housing-related economic stress. The last housing indicator regarded mortgage and rent expenditures. Respondents were queried about monthly first mortgage, second mortgage and rent payments. An indicator of economic stress was calculated by summing these payments and multiplying by 12 to arrive at an annual housing expenditure. The annual housing expenditure was divided by annual family income to determine the share of annual income spent on housing. If the share of income spent on housing was greater than 30 percent, an amount commonly considered to represent a household that is housing cost burdened (US/HUD, 2007), we considered this to be an indicator of economic stress. The housing stressors index ranged in value from 0 to 4.

Financial stressors. Economic stressors related to finances were measured via ten variables. Respondents were asked a series of questions regarding possible financial problems they may have experienced in the past 12 months. Specifically, respondents were asked: Have you done any of the following or have any of the following happened as a result of economic problems in the last 12 months?: (1) Sold possessions or cashed in life insurance; (2) Postponed major purchases; (3) Postponed medical care; (4) Borrowed money from friends or relatives; (5) Filed for or taken bankruptcy; (6) Fallen behind in paying bills; (7) Had a creditor call or come to see you to demand payment; (8) Had your wages attached or garnished by a creditor; (9) Had a lien filed against your property because you could not pay a bill; and (10) Had your home, car or other property repossessed. For each question, a 'yes' response was considered an indicator of economic stress. The financial stressors index ranged in value from 0 to 10.

Neighborhood stressors. Economic stressors related to the household's neighborhood were measured via two variables. Respondents were asked to rate their neighborhood as a place to raise children on a five-point scale ranging from 'excellent' to 'poor.' Families residing in neighborhoods self-assessed as being a 'fair' or 'poor' place to raise children were considered to be experiencing neighborhood-related economic stress. Respondents also were asked how safe it was to walk around alone in their neighborhood after dark. Responses ranged from 'completely safe' to 'extremely dangerous.' Families residing in neighborhoods self-assessed as being 'somewhat dangerous' or 'extremely dangerous' also were considered to be experiencing neighborhood-related economic stress. The neighborhood stressors index ranged in value from 0 to 2.

One-third (31.20%) of the respondents reported they had experienced some aspect of housing stress (had an index value > 0) at one or both CDS interviews. Similarly, about one-third (33.57%) reported experiencing some facet of neighborhood stress during the study. It was more common to experience an aspect of financial stress. Three-fourths (74.43%) of the respondents indicated they had experienced an element of financial stress during one or both of the CDS interviews. As seen in Table 1, the mean values for each of the three stressor indices decreased from CDS-I to CDS-II for all children. Housing fell from 0.245 (out of 4) to 0.196; financial decreased from 1.451 (out of 10) to 1.182; neighborhood declined from 0.299 (out of 2) to 0.295. Index values declined across age groups as well, except for neighborhood stressors increasing between interviews from 0.298 to 0.301 for younger children. As seen in Table 2, some children experienced change in these measures. The most change occurred with financial stressors with over 60 percent of the sample experiencing a change. In contrast, about one-fourth of all children in the analytic sample experienced a change in their level of housing or neighborhood stressors.

3.2.3 Time varying covariates

Two time varying covariates were included in these analyses. These were the number of persons in the household and annual total family income. As seen in Table 1, there were about four members on average in each household. The slight mean decrease in household size between waves stemmed from three times as many of the households of older children experiencing a decrease in size versus an increase in size (38.1% and 13.7%, respectively). Mean annual income increased from $50,565 at CDS-I to $70,725 at CDS-II with approximately three-fourths of the sample gaining income between waves.

4. Results

Results from our estimation of equation (1) on the samples of all children, younger children and older children are displayed in Table 3. The columns in each table reflect results for the full analytic sample (columns 1-2), the younger children (columns 3-4) and the older children (columns 5-6) for the various specifications of weight status (OB). In columns (1), (3) and (5), we set $\alpha=0$ in equation (1) (i.e., the binary measure of obesity) and estimated logit fixed effects models. In columns (2), (4) and (6), we set $\alpha=1$ (i.e., the obesity gap). We estimated linear regression fixed effects models for the obesity gap analyses.

	All Children 2 to 14 yrs		Younger Children 2 to 8 yrs		Older Children 8 to 14 yrs	
	Obese	Obesity Gap	Obese	Obesity Gap	Obese	Obesity Gap
	(1)	(2)	(3)	(4)	(5)	(6)
Stressor Indices						
Housing stressors	0.4618** (0.1999)	0.0407*** (0.0144)	0.4308* (0.2213)	0.0563** (0.0218)	0.8376* (0.5021)	0.0171 (0.0173)
Financial stressors	-0.0165 (0.0786)	-0.0051 (0.0057)	-0.0929 (0.0902)	-0.0069 (0.0089)	-0.1842 (0.1633)	-0.0023 (0.0066)
Neighborhood stressors	-0.0058 (0.1819)	-0.0042 (0.0143)	0.0557 (0.2115)	0.0118 (0.0232)	-0.2718 (0.3730)	-0.0189 (0.0158)
Control Variables						
Number of household members	-0.1442 (0.1041)	-0.0049 (0.0081)	0.0198 (0.1295)	0.0039 (0.0131)	-0.4937** (0.1991)	-0.0131 (0.0091)
Annual total family income ($1000)	-0.0022 (0.0023)	-0.0001 (0.0001)	-0.0025 (0.0034)	-0.0003 (0.0003)	-0.0022 (0.0033)	-0.0001 (0.0001)
Constant		0.1500*** (0.0355)		0.1480** (0.0582)		0.1550*** (0.0389)
Number of observations/ respondents	2526 / 1263		1354 / 677		1172 / 586	

Estimated coefficients with standard errors in parentheses. Obese (BMI ≥ 95th percentile). Superscripts of *, **, and *** indicate that the p-value of the coefficient is less than 0.10, 0.05, or 0.01, respectively.

Table 3. Effects of stressor indices and time varying covariates on weight status and gap measures.

The housing stressors index was significantly, positively related to the likelihood of being obese across both specifications of weight status for all children. With respect to the magnitude of the relationship, a one unit increase from the average value of the housing stressors index across both interview waves (from 0.221 to 1.221) led to a 10.7 percentage point increase in the probability of being obese from a mean probability of 20% at CDS-I, and 32.6% increase in the depth (obesity gap) of obesity for those who were obese. (Derived from Tables 1 and 3.) We found no evidence of a relationship between the other forms of economic stressors –financial and neighborhood stressors – and child weight status.

	Obese (1)	Obesity Gap (2)	Obese (3)	Obesity Gap (4)
	Younger Children: 2 to 8 yrs			
	Females		*Males*	
Housing stressors	0.8029** (0.3983)	0.0810*** (0.0279)	0.1895 (0.2949)	0.0245 (0.0346)
Financial stressors	-0.1547 (0.1534)	-0.0038 (0.0125)	-0.0559 (0.1154)	-0.0094 (0.0128)
Neighborhood stressors	0.0034 (0.3349)	0.0055 (0.0351)	0.0787 (0.2899)	0.0105 (0.0314)
Number of observations/ respondents	628 / 314		726 / 363	
	Older Children: 8 to 14 yrs			
	Females		*Males*	
Housing stressors	20.0090 (4476.3590)	0.0231 (0.0217)	0.3925 (0.5196)	0.0120 (0.0271)
Financial stressors	-0.3554 (0.2932)	-0.0053 (0.0077)	-0.1170 (0.2217)	0.0008 (0.0112)
Neighborhood stressors	-1.1994 (0.7814)	-0.0392** (0.0194)	0.2019 (0.5651)	0.0008 (0.0251)
Number of observations/ respondents	590 / 295		582 / 291	

Estimated coefficients with standard errors in parentheses. Obese (BMI ≥ 95th percentile). Superscripts of *, **, and *** indicate that the p-value of the coefficient is less than 0.10, 0.05, or 0.01, respectively. The coefficients on the time varying covariates are suppressed for brevity.

Table 4. Effects of stressor indices and time varying covariates on weight status and gap measures, by age group and gender.

The relationships between economic stressors and child weight status were considered for younger (age 2 to 8 years at the CDS-I interview) and older (age 8 to 14 years) children separately and are reported in Table 3, as well. For younger children, exposure to housing stressors was positively and significantly related to obesity. Examining the magnitude of the relationships as above, a one unit increase in the housing stressors index led to a 10.7 percentage point increase in the probability of being obese. The percent increase for the obesity gap measure was 36.9%. While the percentage point increases in the probability of being obese were the same for all children and only the younger children, the magnitude of the association for the depth of obesity for those who were obese was larger for the sample of younger children than for all children. Unlike younger children, for older children the housing stressors index was significantly positively associated with weight status for only the binary obesity measure. This result indicated that a one unit increase in the housing stressor index was associated with a 9.5 percentage point increase in the likelihood of being obese for older children. The other two measures of economic stress were statistically insignificant for both age groups.

Table 4 provides additional depth to the age-based analyses by examining the relationships between the three economic stressors and weight status by the gender of the child. These analyses revealed that the significant relationship between housing stressors and weight status for younger children was confined to females. For younger females, a one unit increase in the housing stressors index increased the likelihood that a younger female was obese by 19.3 percentage points. The percentage increases in the depth (obesity gap) of obesity for those young females who were obese stemming from a unit increase in the housing stressor index was 57.9. The relationship between housing stressors and weight status for younger males was statistically insignificant. The remaining results were generally consistent with those reported in Table 3. Financial and neighborhood stressors were not related to weight status for younger females or males. For older females and males, none of the stressor indices were related to weight status except for exposure to neighborhood stressors unexpectedly being negatively associated with weight status for older females in the obesity gap examinations.

5. Conclusions

Although efforts to prevent and treat obesity have traditionally emphasized physical activity and diet, it is now becoming more widely recognized that the causes of childhood obesity are complex and multifactorial (Dietz & Robinson, 2008; Eisenmann, 2006). In this paper, we considered the impacts of economic stressors on childhood obesity. Employing an ecological theoretical framework (Bronfenbrenner & Morris, 1998), we considered factors within the microsystem and mesosystem associated with childhood obesity. Specifically, we examined the relationships between economic stress and obesity during childhood and adolescence. In contrast to previous work on this topic, we explicitly controlled for the effects of time invariant unobserved factors (e.g., a child's ability to cope with adverse conditions) which may be correlated with stress through the estimation of fixed effects models. Additionally, we included measures that portray the incidence of obesity. After controlling for these factors, we found statistically significant effects of economic stress on childhood obesity. Consistent with others (e.g., Garasky et al. (2009)), we found that the relationships between economic stressors and child weight status differed across age groups. More specifically, experiencing housing stress was significantly positively associated with a younger girl's probability of being obese, and the depth of her obesity.

The differing associations between stressors and child weight by age groups can be considered within either a developmental family process framework or a developmental neuropsychoendocrine perspective (i.e., maturation of the stress systems). Younger children are more dependent on their caregivers. Middle childhood youth and early adolescents (ages 8 to 14 years in this study) may have additional/alternative sources of support such as peers, coaches, teachers and romantic partners to whom they may turn in times of stress. Additionally, we conjecture that housing stressors may be more apparent to younger children than other stressors. Compared to the other sources of economic stress considered here (financial and neighborhood stressors), housing-related stress may have the most serious ramifications in terms of altering a child's day-to-day routine. There is evidence, for example, showing that frequent moves are negatively related to the well-being of children (Astone & McLanahan, 1994; Tucker et al., 1998). Children and families experiencing the types of housing stress examined here may have less control over their food choices and physical activities. Girls in particular are less physically active (Sallis et al., 2000), are more susceptible to stress (Rudolph, 2002), and are more likely to use food as a coping mechanism (Larsen et al., 2006.)

We offer five recommendations for future research based on the results of this study. First, other indicators of economic stress should be considered. For example, families often contend with finding adequate and affordable child care which may impact the health and well being of their children, especially their younger children. Second, although examining the mechanisms underlying the relationship between economic stressors and obesity was outside the scope of this study, several potential pathways could be examined. One possibility is parenting. Several studies suggest that economic pressure is associated with lower marital quality, lower parenting quality, and higher levels of depression in children's caretakers, each of which is associated with worse child outcomes (e.g., Kalil & Dunifon, 2007; Robila & Krishnakumar, 2006). Third, our models implicitly assumed that exposure to economic stressors affected child weight status, but that child weight status did not affect the economic stress experienced in the child's household. While the consensus in the literature is that our models reflect the appropriate direction of the relationship when considering economic stressors, future research may want to consider whether the direction of the relationship is reversed when examining non-economic stressors. For example, parents with overweight children may become stressed due to the stigmatization and discrimination associated with weight in our society (e,g., Puhl & Brownell, 2006) or may experience stress as a result of negative perceptions of their appearance (Kraig & Keel, 2001). Fourth, researchers may wish to investigate the effects of longer-term economic shifts within families on child weight status using the longitudinal study design employed here. For example, the effects of income volatility, changes in family structure, and job loss could be considered. Similarly, longer term exposure to stressors could be examined. Lastly, while this article has not examined the effect of the social safety net on childhood obesity, future research may wish to consider how the numerous assistance programs in the U.S. interact to help families mitigate economic challenges and stress.

Dietz and Robinson (2008) contend that it is unlikely that the problem of obesity will respond to a single intervention. A clear policy implication from this research is the identification of a new avenue for reducing the likelihood that a child will be obese. Based on our findings, reductions in housing stress will lead to reductions in the extent and depth of childhood obesity. From a policy perspective, efforts to address housing stress such as the provision of rent vouchers and current programs targeting families facing foreclosure may have a potential added benefit of reducing childhood obesity and associated health care costs to U.S. government programs such as Medicaid and State Child Health Insurance Programs.

6. Acknowledgment

We wish to thank Yemisi Kuku for tremendous research assistance.

7. References

Allison, P. (2005). *Fixed Effects Regression Methods for Longitudinal Data Using SAS.* SAS Institute: Cary, NC.

Anderson, P., & Butcher, K. (2006). Childhood obesity: trends and potential causes. *The Future of Children, 16,* 19-45.

Anderson, P., Butcher, K., & Levine, P. (2003). Economic perspectives on childhood obesity. *Economic Perspectives, 27,* 30-48.

Aneshensel, C. (1992). Social stress: Theory and research. *Annual Review of Sociology, 18*, 15-38.

Astone, N., & McLanahan, S. (1994). Family structure, residential mobility, and school dropout: A research note. *Demography, 31,* 575-584.

Barlow, S. (2007). Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report. *Pediatrics, 120,* S164-S192.

Bjorntorp, P. (2001). Do stress reactions cause abdominal obesity and comorbidities? *Obesity Reviews, 2,* 73-86.

Boss, P. (1998). *Family Stress Management*. Newbury Park, CA: Sage.

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. In R. M. Lerner (Ed.), *Handbook of child psychology theory* (5th ed., Vol. 1, pp. 993-1028). New York: Wiley.

Brooks, T., Harris, S., Thrall, J., & Woods, E. (2002). Association of adolescent risk behaviors with mental health symptoms in high school students. *Journal of Adolescent Health, 31,* 240-246.

Chrousos, G., & Gold, P. (1992). The concepts of stress and stress system disorders: Overview of physical and behavioral homeostasis. *Journal of the American Medical Association, 267*, 1244-52.

Crossman, A., Deborah, A., & Benin, M. (2006). The family environment and American adolescents' risk of obesity as young adults. *Social Science & Medicine, 63,* 2255-2267.

Cutler, D., Glaeser, E., & Shapiro, J. (2003). Why have Americans become more obese? *Journal of Economic Perspectives, 17,* 93-118.

Dearing, E., McCartney, K., & Taylor, B. (2006). Within-child associations between family income and externalizing and internalizing problems. *Developmental Psychology, 42,* 237-252.

Dietz, W., & Robinson, T. (2008). What can we do to control childhood obesity? *The Annals of the American Academy of Political and Social Science, 615,* 222-224.

Dowda, M., Ainsworth, B., Addy, C., Saunders, R., & Riner, W. (2001). Environmental influences, physical activity, and weight status in 8- to 16-year olds. *Archives of Pediatrics & Adolescent Medicine, 155,* 711-717.

Eisenmann, J. (2006). Insight into the causes of the recent secular trend in pediatric obesity: common sense does not always prevail for complex, multi-factorial phenotypes. *Preventive Medicine, 42,* 329-335.

Evans, G., & English, K. (2002). The environment of poverty: multiple stressor exposure, psychphysiological stress, and socioemotional adjustment. *Child Development, 73,* 1238-1248.

Fletcher, C., Garasky, S., & Nielsen, R. (2005). Transportation hardship: are you better off with a car? *Journal of Family and Economic Issues, 26,* 323-343.

Freedman, D., Mei, Z., Srinivasan, S., Berenson, G., & Dietz, W. (2007). Cardiovascular risk factors and excess adiposity among overweight children and adolescents: the Bogalusa Heart Study. *Journal of Pediatrics, 150,* 12–17.

Frydenberg, E., & Lewis, R. (2004). Adolescent coping: the different ways in which boys and girls cope. *Journal of Adolescence, 14,* 119-133.

Garasky, S., Stewart, S., Lohman, B., Gundersen, C., & Eisenmann, J. (2009). Family stressors and childhood obesity. *Social Science Research, 38,* 755-766.

Gee, G., & Walsemann, K. (2009). Does health predict the reporting of racial discrimination or do reports of discrimination predict health? findings from the National Longitudinal Study of Youth. *Social Science & Medicine, 68,* 1676-1684.

Gershoff, E., Aber, J., Raver, C., & Lennon, M. (2007). Income is not enough: Incorporating material hardship into models of income associations with parenting and child development. *Child Development*, *78*, 70-95.

Gibson, L., Byrne, S., Davis, E., Blair, E., Jacoby, P., & Zubrick, S. (2007). The role of family and maternal factors in childhood obesity. *Medical Journal of Australia, 186,* 591-595.

Goodman, E. (1999). The role of socioeconomic status gradients in explaining differences in US adolescents' health. *American Journal of Public Health, 89,* 1522-1528.

Gottholmseder, G., Nowotny, K., Pruckner, G., & Theurl, E. (2009). Stress perception and commuting. *Health Economics, 18,* 559–576.

Gundersen, C., Lohman, B., Garasky, S., Stewart, S., & Eisenmann, J. (2008). Food security, maternal stressors, and overweight among low-income U.S. children: results from National Health and Nutrition Examination Survey (1999-2002). *Pediatrics, 122,* e529-e540.

Hepworth, J. (2004). The social construction of eating disorders. In J. Germov & L. Williams (Eds.), *A Sociology of Food and Nutrition 2nd ed.* (pp. 368-382). Oxford University Press: South Melbourne, Australia.

Jolliffe, D. (2004). Continuous and robust measures of the overweight epidemic: 1971-2000. *Demography, 41,* 303-314.

Kalil, A., & Dunifon, R. (2007). Maternal work and welfare use and child well-being: evidence from 6 years of data from the Women's Employment Study. *Children and Youth Services Review, 29,* 742-761.

Kim, J., & Garman, E. (2003). Financial stress and absenteeism: an empirically derived research model. *Financial Counseling and Planning, 14,* 31-42.

Kim, J., Sorhaindo, B., & Garman, E. (2006). Relationship between financial stress and workplace absenteeism of credit counseling clients. *Journal of Family and Economic Issues, 27,* 458-478.

Koch, F., Sepa, A., & Ludvigsson, J. (2008). Psychological stress and obesity. *Journal of Pediatrics, 153,* 839-844.

Kort-Butler, L. (2009). Coping styles and sex differences in depressive symptoms and delinquent behavior. *Journal of Youth and Adolescence, 38,* 122-136.

Kraig, K., & Keel, P. (2001). Weight-based stigmatization in children. *International Journal of Obesity, 25,* 1661-1666.

Krause, E., Mendelson, T., & Lynch, T. (2003). Childhood emotional invalidation and adult psychological distress: the mediating role of emotional inhibition. *Child Abuse & Neglect, 27,* 199-213.

Kuczmarski, R., Ogden, C., Guo, S., Grummer-Strawn, L., Flegal, K., Mei, Z., Wei, R., Curtin, L., Roche, A., & Johnson, C. (2002). 2000 CDC growth charts for the United States: methods and development. *Vital Health Statistics, 11,* 1-190.

Kumanyika, S., & Grier, S. (2006). Targeting interventions for ethnic minority and low-income populations. *The Future of Children, 16,* 187-207.

Larsen, J., van Strien, T., Eisinga, R., & Engles, R. (2006). Gender differences in the association between alexithymia and emotional eating in obese individuals. *Journal of Psychosomatic Research, 60,* 237-243.

Larson, K., Russ, S., Crall, J., & Halfon, N. (2008). Influence of multiple social risks on children's health. *Pediatrics, 121,* 337-344.

Lohman, B., Stewart, S., Gundersen, C., Garasky, S., & Eisenmann, J. (2009). Adolescent overweight and obesity: links to food insecurity and individual, maternal, and family stressors. *Journal of Adolescent Health, 45,* 230-237

Loureiro, M., & Nayga, R. (2005). International dimensions of obesity and overweight related problems: an economic perspective. *American Journal of Agricultural Economics, 87,* 1147-1153.

MacFadyen, A., MacFadyen, H., & Prince, N. (1996). Economic stress and psychological well-being: an economic psychology framework. *Journal of Economic Psychology, 17,* 291-311.

Marder, W., & Chang, S. (2006). *Childhood Obesity: Costs, Treatment Patterns, Disparities in Care, and Prevalent Medical Conditions*. Thomson: Stamford, CT.

Marniemi, J., Kronholm, E., Aunola, S., Toikka, T., Mattlar, C., Koskenvuo, M., & Rönnemaa, T. (2002). Visceral fat and psychosocial stress in identical twins discordant for obesity. *Journal of Internal Medicine, 251,* 35-43.

McGuigan, F. (1999). *Encyclopedia of Stress.* Allyn and Bacon: Boston.

Mikolajczyk, R., Ansari, W., & Maxwell, A. (2009). Food consumption frequency and perceived stress and depressive symptoms among students in three European countries. *Nutrition Journal, 8,* 31-39.

Ogden, C., Carroll, M., Curtin, L., Lamb, M., & Flegal, K. (2010). Prevalence of high body mass index in US children and adolescents, 2007-2008. *Journal of the American Medical Association, 303,* 242-249.

Ogden, C., Kuczmarski, R., Flegal, K., Mei, Z., Guo, S., Wei, R., Grummer-Strawn, L., Curtin, L., Roche, A., & Johnson, C. (2002). Centers for Disease Control and Prevention 2000 growth charts for the United States: improvements to the 1977 National Center for Health Statistics Version. *Pediatrics, 109,* 45-60.

Pearlin, L., Menaghan, E., Lieberman, M., & Mullan, J. (1981). The stress process. *Journal of Health and Social Behavior, 22,* 337-356.

Pearlin, L., Schieman, S., Fazio, E., & Meersman, S. (2005). Stress, health, and the life course: Some conceptual perspectives. *Journal of Health and Social Behavior, 46,* 205-219.

PSID (Panel Study of Income Dynamics). 2005. *An Overview of the Panel Study of Income Dynamics.* http://psidonline.isr.umich.edu/Guide/Overview.html [13 May 2005].

PSID (Panel Study of Income Dynamics). 2008a. *The PSID-CDS Design at-a-Glance.* http://psidonline.isr.umich.edu/CDS/questionnaires/cdsover.pdf [21 September 2008].

PSID (Panel Study of Income Dynamics). 2008b. *What is the CDS?* http://psidonline.isr.umich.edu/CDS/researchdesign.html [21 September 2008].

Puhl, R., & Brownell, K. (2006). Confronting and coping with weight stigma: an investigation of overweight and obese adults. *Obesity, 14,* 1802-1815.

Puhl, R., & Latner, J. (2007). Stigma, obesity, and the health of the nation's children. *Psychological Bulletin, 133,* 557-580.

Raman, R. (2002). Obesity and health risks. *Journal of the American College of Nutrition, 21,* 134S-139S.

Robila, M., & Krishnakumar, A. (2006). Economic pressure and children's psychological functioning. *Journal of Child and Family Studies, 15,* 435-443.

Rudolph, K. (2002). Gender differences in emotional responses to interpersonal stress during adolescence. *Journal of Adolescent Health, 30,* 3-13.

Sallis, J., Prochaska, J., & Taylor, W. (2000). A review of correlates of physical activity of children and adolescents. *Medicine & Science in Sports & Exercise, 32,* 963-975.

Sampson, R., Raudenbush, S., & Earis, F. (1997). Neighborhoods and violent crimes: A multilevel study of collective efficacy. *Science*, *277*, 918-924.

Schilling, E., Aseltine, R., & Gore, S. (2008). The impact of cumulative childhood adversity on young adult mental health: measures, models, and interpretations. *Social Science & Medicine, 66,* 1140-1151.

Siervo, M., Wells, J., & Cizza, G. (2009). The contribution of psychosocial stress to the obesity epidemic: an evolutionary approach. Hormone and Metabolic Research, *41,* 261-270.

Skinner, A., Mayer, M., Flower, K., & Weinberger, M. (2008). Health status and health care expenditures in a nationally representative sample: how do overweight and healthy-weight children compare? *Pediatrics, 121,* e269-e277.

Spencer, M. (2001). Resiliency and fragility factors associated with the contextual experiences of low-resource urban African-American male youth and families. In *Does it Take a Village? Community Effects on Children, Adolescents, and Families*, edited by A. Booth and A.C. Crouter. Mahwah, N.J.: Lawrence Erlbaum Associates, Inc.

Strauss, R. (1999). Childhood obesity. *Current Problems in Pediatrics, 29,* 5-29.

Tucker, C., Marx, J., & Long L. (1998). Moving on: resident mobility and children's school lives. *Sociology of Education, 71,* 111-129

US/DHHS (U.S. Department of Health and Human Services). (2000). *Healthy People 2010, 2nd ed; With Understanding and Improving Health and Objectives for Improving Health.* 2 vols. U.S. Government Printing Office: Washington, DC.

US/HUD (U.S. Department of Housing and Urban Development). (2007). *Affordable Housing.* http://www.hud.gov/offices/cpd/affordablehousing/index.cfm [16 October 2007].

US/Office of the President. (2010). *Solving the Problem of Childhood Obesity within a Generation: Report to the President*. White House Task Force on Childhood Obesity: Washington, DC.

van Jaarsveld, C., Fidler, J., Steptoe, A., Boniface, D., & Wardle, J. (2009). Perceived stress and weight gain in adolescence: a longitudinal analysis. *Obesity, 17,* 2155-2161.

Wadsworth, M., & Compas, B. (2002). Coping with family conflict and economic strain: the adolescent perspective. *Journal of Research on Adolescence, 12,* 243-274.

Wang, Y., & Zhang, Q. (2006). Are American children and adolescents of low socioeconomic status at increased risk of obesity? Changes in the association between overweight and family income between 1971 and 2002. *The American Journal of Clinical Nutrition, 84,* 707-716.

White, J., & Klein, D. (2002). *Family Theories,* 2nd ed. Sage: Thousand Oaks, CA.

Wickrama, K., Elder, G. Jr., & Abraham, T. (2007). Rurality and ethnicity in adolescent physical illness: Are children of the growing rural Latino population at excess health risk? *Journal of Rural Health, 23*, 228-237.

7

Critical Appraisal of Selected Body Composition Data Acquisition Techniques in Public Health

Steven Provyn, Aldo Scafoglieri, Jonathan Tresignie,
Céline Lumé, Jan Pieter Clarys and Ivan Bautmans
Vrije Universiteit Brussel
Belgium

1. Introduction

Human body composition (BC) may be the most typical health-related discipline that enables both "easy indexes" and "complicated technology" for the same purpose, e.g. measuring quantities of adipose tissue, muscle, skin, bone and its minerals to predict health risks.

Human BC may also be the most confusing health-related discipline because of the mixture of apparent corresponding and analogue terminology. Fat versus adipose tissue, fat-free mass versus lean-body mass versus adipose-tissue -free mass, visceral fat versus internal fat or abdominal fat. For the non-expert, who is often the clinical user, this is very confusing, especially because experts themselves do not always differentiate as they should between fat and adipose tissue.

Methods for assessing human body composition are applied in many fields. In particular, the assessment or prediction of 'total body fat' is a common, popular and, at the same time, important element of public health, physical anthropology, sport and exercise sciences and, more specifically, of kinanthropometry, physiology, biomechanics, auxology and ergonomics. It is also general knowledge that monitoring adiposity is a dominant factor in analysing body composition, and that skinfold (SF) measurements (and quantities derived from them) play a key role in the prediction of adiposity. In addition, skinfolds have specific applications in occupational biomechanics, human hydrodynamics, drug quantification, diabetes, coronary heart disease, nutrition, endocrinology, hypertension, anorexia nervosa and in many epidemiological and human growth studies. Consequently, the SF is also a central factor in adipose tissue patterning, in 'fat' distribution studies, in somatotyping and in commercialised systems for monitoring adiposity and proportional mass (Edwards, 1951; Garn, 1955; Heath and Carter, 1967; Garn et al., 1971; Duquet et al., 1977; Mueller and Stallones, 1981; Jurimae et a l., 2005; Jurimae et al., 2007; Tafeit et al., 2007).

Given the easy accessibility of the subcutaneous layer and its non-invasive nature, this interest in skinfolds has led to a proliferation of SF applications and formulae. In the literature, there are over 1000 articles dealing directly with SF measurement, both in applied and fundamental research. Altogether, more than 600 equations have been developed to predict body fat from skinfolds (Lohman, 1981; Martin et al., 1985; Clarys et al., 1987; Clarys et al., 1999; Provyn et al. 2010, 2011).

Body fat or adipose tissue prediction formulae can be subdivided into regression equations based on anthropometric variables called anthropometric-based formulae (ABF) and into equations, once again based on anthropometric variables, but initially estimating density from which BF or AT are calculated in a second phase. These formulae will be called density-based formulae (DBF), and thereafter %AT is obtained with the Siri (1956) or Brozek (1963) conversion formulae according to the original publication.

The accuracy of population-specific equations was improved with the addition of breadths and girths in combination with SFs and the use of populations of varied ages and degrees of body adiposity. It is well known, however, that the major weakness of population-specific equations is partly due to their inability to account for ageing and the non-linear relationship between subcutaneous AT and body density (Daniel et al., 2003).

Due to its popularity and oversupply, the choice of a suitable equation is not evident. When selecting the most appropriate equation, using your common sense is advisable and generally accepted, if your selection is based on the characteristics of the population on which the equation was originally validated. In reality its application is not guaranteed. Studies comparing anthropometry with reference body composition techniques conducted on large samples of different ages are necessary (Rolland-Cachera and Brambilla, 2005).

We know that the tissue distribution is different among men and women in particular adipose tissue and muscle tissues. (Clarys et al., 1984; Clarys et al., 1999)

With reference to public health issues, there is a rising demand to improve the performance of techniques and systems while, at the same time, increasing clinical precision. Related research issues require that we consider whether or not we are studying what we think we are and whether or not the measuring techniques we use are consistent. Deeper insight into failures and how to prevent them can be gained by comparing and contrasting the reliability of systems with differing characteristics: electromechanical machinery, bioelectrical analysers, scanners and imaging equipment, etc.

The user, e.g. the researcher, physician, clinician, or therapist depends on this system reliability. In addition to this reliability, other quantities are necessary to ensure field or user reliability of the system. This chapter will give a critical appraisal of common used body composition data acquisition techniques such as, Anthropometry and prediction equations.

2. Anthropometry

Anthropometry refers to the measurement of proportions of the human body in an easy way, and is probably the best-known and most widely used technique for estimating BC both in the laboratory as well as in rural or urban field situations. Beside measuring weight and height, which do not provide any information about a nutritional status, other techniques are used to measure the size and proportion of body segments, (e.g. skinfolds , bone breadth and lengths, circumference and segment depths).

It is known that SF is a central factor in AT patterning and 'fat' distribution studies (Edwards, 1951; Garn et al., 1971; Mueller and Stallones, 1981), in somatotyping (Heath and Carter, 1967), and in various BC issues (Martin et al., 1985). Clarys et al. indicated that its use is not without criticism, herewith referring to compressibility, lack of tissue constancies, skin-thickness-induced errors and not in the least to the examiner's skill level (Clarys et al., 1987; Clarys et al., 2005).

The interest in skinfolds, given the easy accessibility of the subcutaneous layer and its non-invasive nature, has led to a proliferation of SF applications and formulae. Prediction equations have been developed using either linear (population specific) or quadratic (generalised) regression models. Generally speaking, equations have been developed for relatively homogeneous populations and are assumed to be valid only for individuals with similar characteristics, e.g. gender, ethnicity, age or levels of physical activity. Besides this, anthropometric measurements are suitable for measurement error. Their accuracy depends not only on the prediction ability of the used formulae but also on the measurement skills, subject-related factors (e.g. obesity, age, etc.) and the type of calliper (Lohman et al., 1984).

2.1 Quality control and reliability of the panniculus adiposus calliper – A critical appraisal of the all-round skinfold measure

This section describes the state of the art of an on-going critical search of the skinfold based on experimental anatomy over two decades (Clarys et al., 1984; Martin et al., 1985; Martin et al., 1994; Clarys et al., 1999; Marfell-Jones et al., 2003; Martin et al., 2003; Clarys et al., 2005).

The SF-calliper measurement has become a routine laboratory and field method for so many years, that it has obtained the status of 'a tradition'. In other words, this method has become too normal, almost too obvious to be analysed. Hägar (1991) stated that "two important assumptions must be made in the calculation of body SF measurements:

1. Subcutaneous fat constitutes a constant proportion of total body fat over all ranges of body weight; and
2. Measurement sites are representative of all subcutaneous fat."

This statement is, at best, doubtful (Hägar, 1991). What is really being measured is the thickness of a double skinfold and compressed subcutaneous adipose tissue (Figure 1).

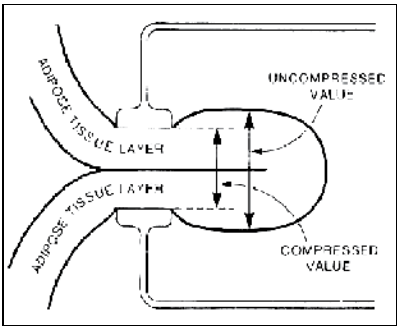


Fig. 1. Schematic representation of a double skinfold and compressed subcutaneous AT with application of a calliper

To infer the mass of fat in the body from this measurement requires another series of assumptions whose validity has never been seriously challenged (Clarys et al., 1987). In order to review the assumptions associated with calliper adiposity transformations, our 'step-by-step' model (Marfell-Jones et al., 2003; Clarys et al., 2005; Clarys et al., 2009) is relevant. The transformation from calliper reading to total body adiposity can be divided into a number of steps. The thickness of a compressed double layer of skin and subcutaneous AT should be representative of the uncompressed single layer of adipose tissue. This should indicate the total subcutaneous adiposity from which internal and whole-body adiposity can be predicted. Based on a pooled cadaver data, Brussels Cadaver Analysis Study (BCAS), (N=51), we have reviewed again the facts, assumptions and hazards to be taken into account in the transformation of SFs to whole-body AT mass.

Figure 2 presents a flow chart of the systematic step-by-step reasoning behind the calliper reading with its associated seven assumptions. Each of these steps and assumptions is considered separately in chronological order:

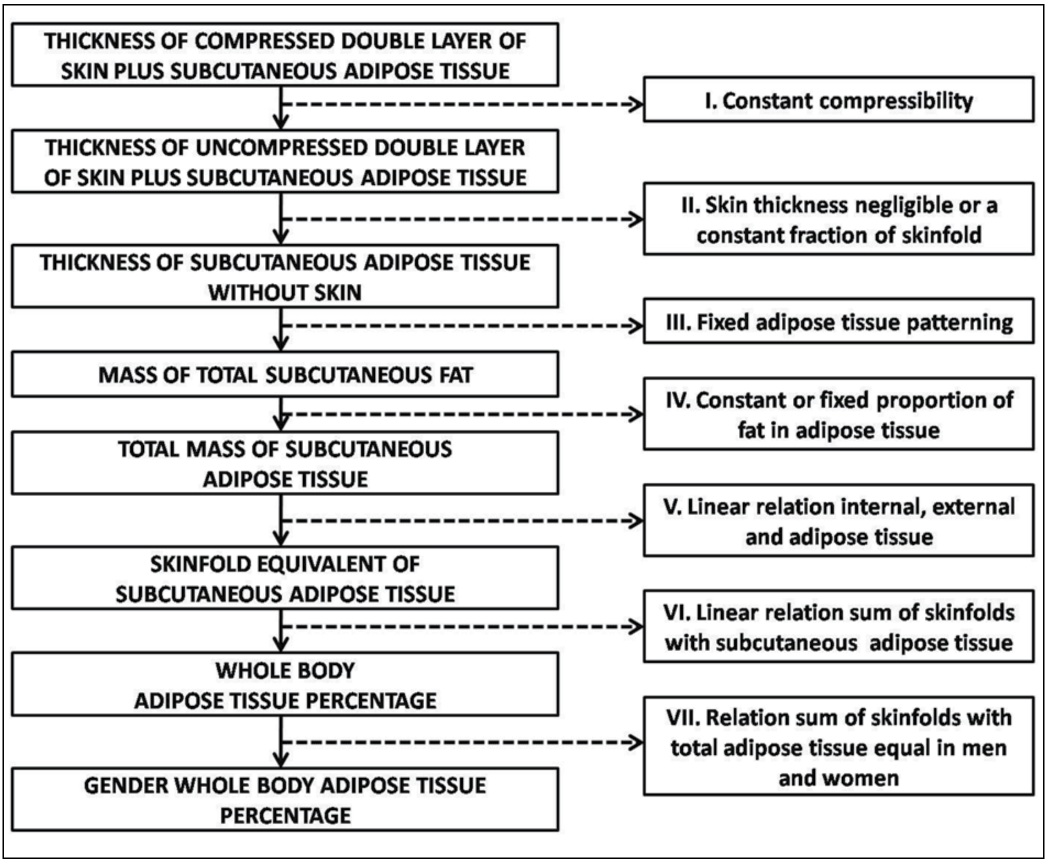


Fig. 2. Flow chart of the transformation from skinfold to total body adiposity; eight possible steps (left) and possible assumptions (right) (Clarys et al., 2009)

2.1.1 Assumption I (constant compressibility)

A calliper produces a constant SF compressibility. Most users of the calliper adopt some strategy to standardise the calliper reading in spite of its dynamic characteristics. Some wait “for all needle movements to cease” before taking the reading, while others record after “an initial rapid phase of the movement” or take the reading after two or four seconds of applying pressure. In addition to the dynamic compressibility, there is also a static element. Even after standardising the timing of the calliper reading, similar thicknesses of AT may yield different calliper values due to different degrees of tissue compressibility. Since the BCAS data include both SF thickness and the direct depth measurement (after incision) of the thickness of the subcutaneous adipose tissue layer, skinfold compressibility could be obtained for each site (Marfell-Jones et al., 2003). However, it was found that SF compressibility is by no means constant.

2.1.2 Assumption II (skin thickness negligible or a constant fraction of skinfold)

Skin thickness is a negligible part or a constant fraction of the skinfold. All SF measurements contain a double layer of skin of unknown thickness. If this is very small in comparison to the SF measurement, its influence may be negligible. Data on skin thickness are sparse. Comprehensive skin thickness and surface data are to be found in various literature (Clarys et al., 1988; Verbraecken et al., 2006; Clarys et al., 2008). The site where the effect of skin thickness was most marked was the subscapular, where skin thickness accounted for 28.1% of the SF reading (34.0% for males, 23.9% for females). The subscapular and triceps sites are most commonly used for predicting whole-body values but have quite different proportions of skin (Clarys et al., 1987; Clarys et al., 2005). Consequently, on the basis of skin thickness, the subscapular skinfold should be a poorer predictor than SFs at arm and leg sites.

2.1.3 Assumption III (fixed adipose tissue patterning)

Adipose tissue patterning is fixed (equal) all over the body. "Fat patterning" refers to differences in the anatomical placement of AT (Mueller, 1985) and therefore should be referred to more accurately as "adipose tissue patterning".

The patterning of subcutaneous AT is known to exhibit very large variations between individuals (Mueller and Stallones, 1981; Clarys et al., 1988; Martin et al., 2003; Clarys et al., 2005).

To assess the value of various sites as predictors of subcutaneous adiposity, correlations have been determined between calliper and incision thickness with the dissected subcutaneous adipose tissue mass (Clarys et al., 1987). An unexpected finding is the high correlation for lower limb sites. Of the six best sites, all but one were on the lower limbs. The triceps, a highly favoured site for 'fat' prediction and considered to be the single indicator of AT (e.g. in digitised commercial devices) ranked a poor eleventh. The best predictors were front thigh, medial calf, rear thigh and supra-spinal. To summarise, under no circumstances is adipose tissue patterning divided equally over the body.

2.1.4 Assumption IV (constant or fixed proportion of fat in adipose tissue)

Predicting human body fat is conditional on the knowledge of the fat content of, or in, AT. Even if the exact mass of subcutaneous adipose tissue is known, the prediction of subcutaneous fat mass requires some assumptions concerning the fat content of AT. Reported values range from 5.2 to 94.1% but are generally in the range 60 - 85%. In addition, the fat content of adipose tissue increases with increasing adiposity. Taking into account these considerations, which are compounded by the fact that 'fat' is ether extractable, whereas 'adipose tissue' is an anatomical – morphological – entity, confusion over the two (which occurs too often) should be avoided by eliminating 'fat' terminology from all morphologically based predictions of adiposity (Clarys et al. 1987).

2.1.5 Assumption V (Linear relation between internal, external and total adipose tissue)

A high correlation between internal, external and total AT obtained from skinfolds is essential. From evidence based on cadaver studies, it is assumed that, both in male and female subjects, any excess of adipose tissue is piled up subcutaneously, intramuscularly and internally, mostly in the trunk. The amount of intramuscular fat in the obese should not be underestimated and should therefore be considered as a third component. However, in the cadaver analysis, the intramuscular amount was allocated to the internal AT. SF callipers are only able to estimate subcutaneous adiposity. In order to estimate total body adiposity, some assumptions must be made about the relationship between internal and subcutaneous (external) adipose tissue. Figures 3 and 4 confirm a (very) good relation between whole body AT and both internal and external AT for men and women.

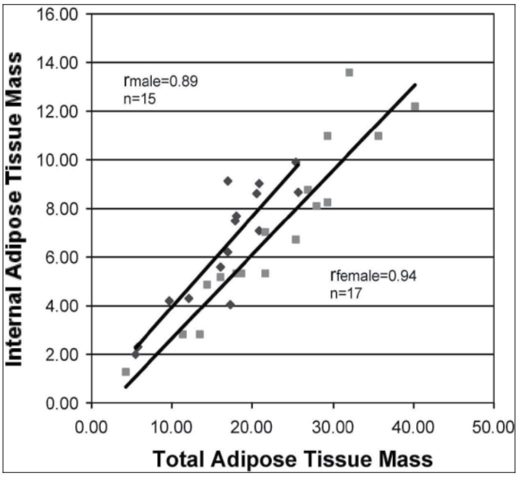


Fig. 3. Total adipose tissue mass versus internal adipose tissue mass (Clarys et al., 2005)

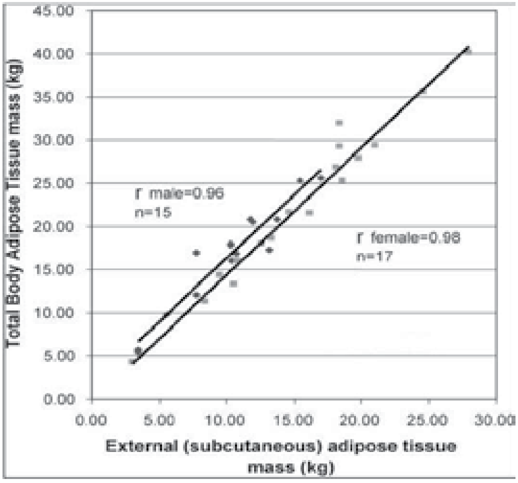


Fig. 4. External (subcutaneous) versus total body adipose tissue (Clarys et al., 2005)

If, in these circumstances, internal adiposity stores are proportional to subcutaneous adiposity, this relationship provides a rationale for use of skinfold callipers. Alternatively, the internal AT may be negligible when compared with subcutaneous adipose tissue, once again providing some justification for the use of callipers. If, however, there is no significant relationship between internal and subcutaneous AT masses, and/or internal adiposity stores are far from negligible, then there cannot be any evidence-based prediction of total body adiposity and, concomitantly, there is no justification for using calliper measurements if these do not correlate with the above. Data indicate a good correlation between external and internal mass in both men (r=0.72) and women (r=0.86) (Figure 5). Almost all the assumptions necessary to convert SF calliper readings to percentages of ether-extractable fat are clearly unfounded, which supports our lack of confidence in the correctness of any whole-body fat prediction that depends on such assumptions. For this reason, once again we recommend the complete rejection of using the term 'fat', in favour of the term 'adipose tissue', which is, in fact, what is actually being measured by SF callipers.

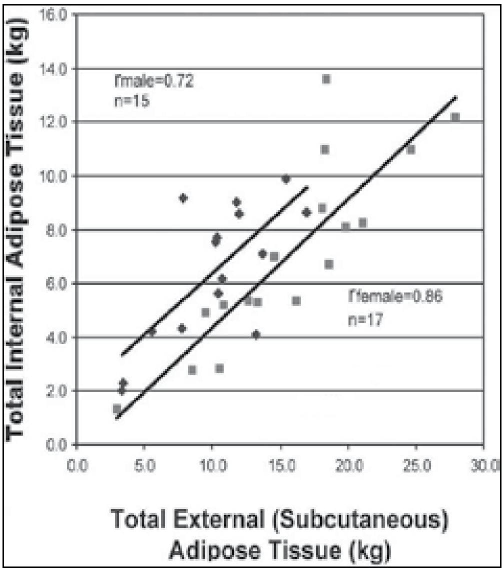


Fig. 5. External (subcutaneous) versus internal (visceral + intermuscular adipose tissue) (Clarys et al., 2005)

2.1.6 Assumption VI (linear relation sum of skinfolds and subcutaneous adipose tissue)

Skinfolds relate to (external) subcutaneous and total AT in men and women.

Having rejected the concept of body fat prediction, we then considered whether total body adiposity could be confidently predicted from skinfolds. To achieve this, SF measurements would need to predict subcutaneous adipose tissue mass adequately, and there would have to be a strong relationship between the latter and total body adiposity. The most commonly used sites for SF measurements (in a variety of combinations) are triceps, subscapular, biceps, iliac crest, supraspinale, abdominal, front thigh and medial calf. The use of all of these sites gives an achievable, reasonably comprehensive coverage of the body's subcutaneous AT deposition.

For this reason, the relationship between the sum of these eight skinfolds and the subcutaneous adipose tissue masses of all those BCAS subjects for whom these data were available (n=20) was examined. Figure 6 shows a strong significant correlation between these entities in men (r=0.82), but a rather poor relationship in women (r=0.56).

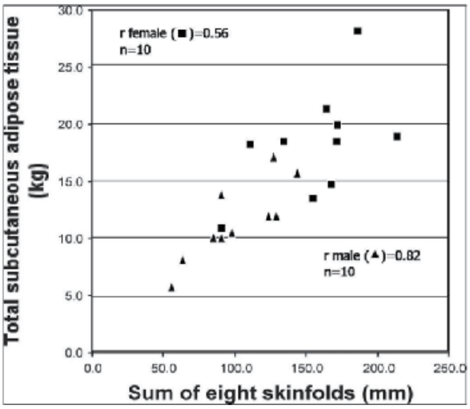


Fig. 6. Sum of skinfolds (8) vs. external (subcutaneous) adipose tissue (Clarys et al., 2005)

2.1.7 Assumption VII (Relation sum of skinfolds and total adipose tissue equal in men and women)

The important difference in basic relations as seen in assumption VI with the results of figure 6 jeopardise the use of SF measurements e.g. SF for men only seems somewhat hazardous. The reason for this gender difference are not be found in the direct relation between external (subcutaneous) AT and whole body AT. Figure 4, for that matter, indicate perfect correlations. If, however, the relation of the sum of skinfolds, for both genders would relate well with the direct whole body AT. The findings as described in assumption VI would be overruled. Unfortunately this is not the case. A r=0.48 in female against r=0.86 in men rejects the idea of equal relations between gender. Because of the difference in BC tissue distribution between men and women (e.g. AT and muscle mass), the use of skinfolds in whatever combination will remain hazardous in women (Clarys et al., 2005; Clarys et al., 2009)

3. Quality control and suitability of age, ethnicity and activity-matched prediction formulae for adiposity

In order to investigate the validity in terms of quality for application in other groups and individuals. Density- and anthropometry-based equations for the determination of %AT have been selected for a sample of white female adults with a broad age-range and different lifestyles. For this purpose, DXA was used as a comparative measure.

3.1 Methodology

One hundred and twenty eight subjects, from the Merseyside, United Kingdom, were recruited through a notice placed on the Liverpool John Moores university website, at health promotion sites used by Liverpool City Council, and at adjacent academic institutions. The subjects' characteristics and profile are shown in Table 1.

	Female			Male		
	Mixed N=54	Actives N=27	Non actives N=27	Mixed N=74	Actives N=55	Non actives N=19
Paramètres	$\bar{x}$ ± SD	$\bar{x}$ ± SD	$\bar{x}$ ± SD	$\bar{x}$ ± SD	$\bar{x}$ ± SD	$\bar{x}$ ± SD
Âge (years)	30.9 ± 8.5	30.4 ± 9.1	31.5 ± 7.9	34.4 ± 14.1	35.4 ± 14.5	31.5 ± 13.1
Height (cm)	164.7 ± 5.8	163.3 ± 5.9	166.1 ± 5.5	178.9 ± 7.1	179.2 ± 6.6	178.1 ± 8.6
Weight (kg)	63.4 ± 11.1	60.7 ± 7.1	66.1 ± 13.6	80.3 ± 9.6	81.1 ± 9.9	78.1 ± 8.5
BMI (kg/m^2)*	23.3 ± 3.4	22.7 ± 1.9	23.9 ± 4.3	23.3 ± 3.4	25.3 ± 3.0	24.6 ± 2.2
Underweight (N)	2	-	2	-	-	-
Waist/Hip ratio	0.70 ± 0.02	-	0.70 ± 0.02	-	-	-
Normal weight (N)	40	24	16	41	30	11
Waist/Hip ratio	0.73 ± 0.04	0.73 ± 0.04	0.74 ± 0.04	0.83±0.04	0.84 ± 0.04	0.81 ± 0.03
Overweight (N)	9	3	6	29	22	7
Waist/Hip ratio	0.79 ± 0.05	0.80 ± 0.06	0.78 ±0.05	0.89±0.05	0.89 ± 0.05	0.87 ±0.06
Obese (N)	3	-	3	4	3	1
Waist/Hip ratio	0.84 ± 0.04	-	0.84 ± 0.04	0.92±0.07	0.89±0.04	1.0
DXA % BF	28.0 ± 6.2	26.1 ± 4.8	29.9 ± 6.9	16.9 ± 4.6	16.7 ± 4.8	17.3 ± 4.3
Hours sport/week	4.5 ±4.3	7.7 ± 3.7	1.2 ± 1.2	6.6 ±4.7	8.2 ± 4.3	1.8 ± 1.2

BF = body fat; BMI= body mass index; DXA= dual energy X-ray absorptiometry; SD= standard deviation; *=BMI classification according to the World Health Organisation: underweight=BMI<18.5; normal weight=18.5<BMI<25; overweight=25≤BMI<30; obese= BMI≥30; $\bar{x}$ = mean

Table 1. Characteristics / profile of the participating subjects;

All subjects received a full written and verbal explanation of the nature of the study before providing informed written consent. Approval for this study was obtained from the University's Research Ethics Committee. Participants presenting osteosynthetic materials (e.g. screws, endoprotheses) or who were pregnant were excluded from the study.

After explaining the measurement procedure, the participants were assessed for anthropometry (number of variables corresponding with the ad hoc %AT constituents) and a DXA scan on the same visit. Participants were asked to refrain from consuming alcohol for 24 hours and food and beverages (except water) for three hours prior to the test session. Participants wore lightweight clothing without zips, buttons or any other metal and removed all jewellery prior to the test protocol. The amount of activity per week was deducted from self-reported information. To separate daily and work-related activity from defined sport participations an arbitrary threshold of four hours was chosen. Less than four hours activity corresponded to "low active" and more than four hours activity per week was considered "high active". The rationale of this choice is a subjective interpretation of daily/weekly activity e.g. undefined walking and house hold activities as against all defined supplementary activity. Subjects were subdivided into three groups: i) high active group (≥4 hour's exercise/week), ii) low active (≤ 4 hours exercise/week) and iii) all subjects (regardless of their physical activity level).

Since anthropometry is an anatomical derivative and because DXA produce value levels that are morphological rather than chemical, all outcomes were labelled as AT instead of BF from here onwards (Clarys et al., 2005; Bolotin, 2007; Provyn et al., 2008).

The research personnel for DXA and anthropometry respectively remained the same throughout the entire examination period. In order to exclude bias due to inter-observer variation, the order in which subjects were assessed was randomised.

3.1.1 Anthropometry

Ninety one (91) formulae corresponded to the inclusion criteria (e.g. gender = female or male, ethnicity = white, age range between 18 and 75 years), among which 34 ABF and 57 DBF fitted within the criteria (Table 2).

Equation reference	Number of equations	Observed population range in years		N	♀/♂	Activity level	Concept of estimation
			$\bar{x}$ + SD				
Wilmore & Behnke 1970	3	18-48	21,4 ± 3,8	128	♀	NM	A
	3	18-48	21,4 ± 3,8	128	♀	NM	D
Katch & Mc Ardle 1973	1	17-24	20,3 ± 1,8	69	♀	Gymnast	D
Durnin & Womersley 1974	15	16-68	-	272	♀	Mixed	D
	15	17-72	-	209	♂	Mixed	D
Pollock et al. 1976	7	18-22	19,7 ± 1,5	95	♂	NM	D
		40-55	44,9 ± 4,8	48	♂	NM	D
Deurenberg et al. 1991	1	16-83	-	708	♀	NM	A
	1	21-66	-	238	♂	NM	A
	1	16-83	-	946	♀/♂	NM	A
Lean et al. 1996	8	18-64	39,9 ± 14,1	84	♀	NM	A
	8	18-64	39,9 ± 14,1	84	♀	NM	D
	8	17-65	40,1 ± 13,1	63	♂	NM	A
	8	17-65	40,1 ± 13,1	63	♂	NM	D
Gallagher et al. 2000	3	14-83*	48,8 ± 17,6	225	♀	NM	A
	3	10-86	48,8 ± 19,2	192	♂	NM	A
	3	10-86	-	417	♀/♂	NM	A
Evans et al.2005	1	18-25	20,6 ± 1,9	102	♀/♂	Athletes	A
	1	18-34	20,9 ± 2,4	30	♀/♂	Athletes	A
	1	18-34	20,8 ± 2,3	132	♀/♂	Athletes	A
Total	91	(= 34 ABF + 57 DBF)					

N=number of subjects for whom the original formulae were developed. NM = not mentioned;
Mixed= active + sedentary; Y = age in years; ABF = Anthropometry-based formulae;
DBF = Density-based formulae

Table 2. Selected adiposity equations applicable on the target population;

With the detailed information of all the necessary variables available, the measurements battery was completed. Measurements were taken from the right side of the body at different locations and included all the sites and measures necessary to calculate the estimation of %AT using all the selected equations.

These anthropometric measurements included:

- **Fourteen skinfolds** : 1) triceps, 2) subscapular, 3) scapular (oblique fold lateral and downward at the inferior angle of the scapula), 4) biceps, 5) forearm 1 (vertical fold at maximum girth, lateral aspect with hand supinated), 6) forearm 2 (vertical fold at max girth, anterior aspect with hand supinated), 7) chest (vertical fold on the midaxillary line at the level of the xiphoid process), 8) iliac crest horizontal, 9) iliac crest vertical, 10) supraspinale, 11) abdominal, 12) front thigh, 13) thigh (vertical fold on the anterior aspect midway between superior aspect of the patella and the anterior superior iliac spine), 14) medial calf.
- **Fourteen girths**: 1) neck, 2) arm girth relaxed, 3) arm girth flexed and tensed, 4) forearm, 5) wrist, 6) chest, 7) waist, 8) waist 1 (midway between inferior margin of the last rib and the crest of the ilium in a horizontal plane and around the pelvis), 9) waist 2 (while standing at umbilical level), 10) gluteal, 11) mid-thigh, 12) knee, 13) calf, 14) ankle.
- **Four breadths**: 1) bi-epicondylar humerus, 2) bi-styloid ulna/radius, 3) bi-epicondylar femur, 4) ankle bi-malleolar.

All anthropometric measurements were taken exactly as described by the respective authors of the selected formulae. If no repeatable descriptions were mentioned, the protocol as described by Martin and Saller (1957), with instructions from The International Society for the Advancement of Kinanthropometry (ISAK)(Norton et al., 1996) was followed.

Anthropometric measurements were taken by a qualified criterion anthropometrist (ISAK Level IV). Body mass was measured to the nearest 0.05 kg with a digital scale (SECA 220, seca gmbh & co, Hamburg, Germany) and stretched stature was measured to the nearest 0.1 cm using a stadiometer with the head in Frankfort plane.

Skinfold thicknesses were measured with a Harpenden calliper (Harpenden skinfold calliper, Baty international, West Sussex, England), girths with a flexible Lufkin steel anthropometric tape (Lufkin W606PM, cooper industries, Ohio, United States) and breadths with a small sliding calliper (Rosscraft Campbell 10 small bone Calliper, Rosscraft Inc., Surrey, Canada) to the nearest 0.1 cm.

Each measurement was taken twice and the mean was calculated. If the difference between the first and the second measure was >5% for skinfold, >1% for girths or breadths, a third measurement was taken and the mean of the two nearest measurements was calculated as the final value.

3.1.2 Dual energy X-ray absorptiometry

The percentage AT is part of the data acquisition from the whole-body DXA scan. Although the data protocol mentions BF, its resulting value level corresponds to anatomical AT values (Nagy and Clair, 2000; Snijder et al., 2002; Bolotin, 2007). All scans were measured according to standard operating procedures using a fan beam dual energy X-ray absorptiometry scanner (Hologic QDR series Delphi A, Bedford, Massachusetts). The scans were analysed using system Hologic QDR software for Windows version 11.2. (© 1986-2001 Hologic inc.). The literature has reported a coefficient of variation for %AT from 1.8 to 6.9% (Madsen et al., 1997; Bachrach, 2000; Nagy and Clair, 2000; Wallace et al., 2005). Each whole-body scan lasted approximately four minutes. All scans were performed by the same experienced examiner who was blinded for the anthropometrical measurements. The DXA was calibrated daily using the anthropometric spine phantom supplied by the manufactures to assess the stability of the measurements. The DXA was also calibrated weekly for body composition using a step phantom (Hologic QDR series Delphi A, Bedford, Massachusetts), that claims the correction of errors related to skin thickness (beam hardening).

3.1.3 Statistical analysis

The percentage of adiposity is part of the data acquisition of the whole body DXA scan. Statistical analysis was conducted using SPSS 19.0 for Windows (© SPSS Inc., Chicago, IL). All variables showed normal distribution (Kolmogorov Smirnov Goodness of Fit test; $p>0.05$). For each equation, the estimated %AT was calculated according to the population for which the formula was developed in the first place and according to age and hours of sports activity per week. The %AT estimated by equations were compared with the %AT as assessed by DXA, using Pearson correlation coefficients and paired sample t-tests. A correlation coefficient higher or equal to 0.70 was chosen as a cut-off value. Significance was set a priori at $p<0.05$. Agreement between %AT using different methods with a correlation in concordance with the previous settings was determined by means of Bland-Altman plots.

3.2 Results

Prediction equations to estimate %AT have been described for different populations according to age and sports participation. To meet with these conditions, the subjects studied were divided into three groups high active (HA), low active (LA) and the combined group (CG), based on the self-reported hours of sport activity per week. For illustrative purposes only we have subdivided our participants into BMI categories, according to the WHO classifications, and according to their waist/hip ratio (Table 1). The percent of AT was calculated for the 91 analysed formulae. Three equations only displayed no significantly different ($p>0.05$) results with DXA (Table 3).

Author	Formula
Wilmore and Behnke 1970	LBW = 8.629 + 0.680 weight – 0.163 subscapular SF – 0.100 triceps SF – 0.054 thigh SF
Pollock et al. 1976	D = 1.07660 – 0.00098 pectoral SF – 0.00053 chest SF
Lean et al. 1996	%BF (AT) = 1.31 triceps SF + 0.430 age – 9.16

LBW=Lean Body Weight; SF=Skinfold; %BF= percentage Body Fat.

Table 3. Remaining equations predicting percentage adipose tissue not significantly different from DXA.

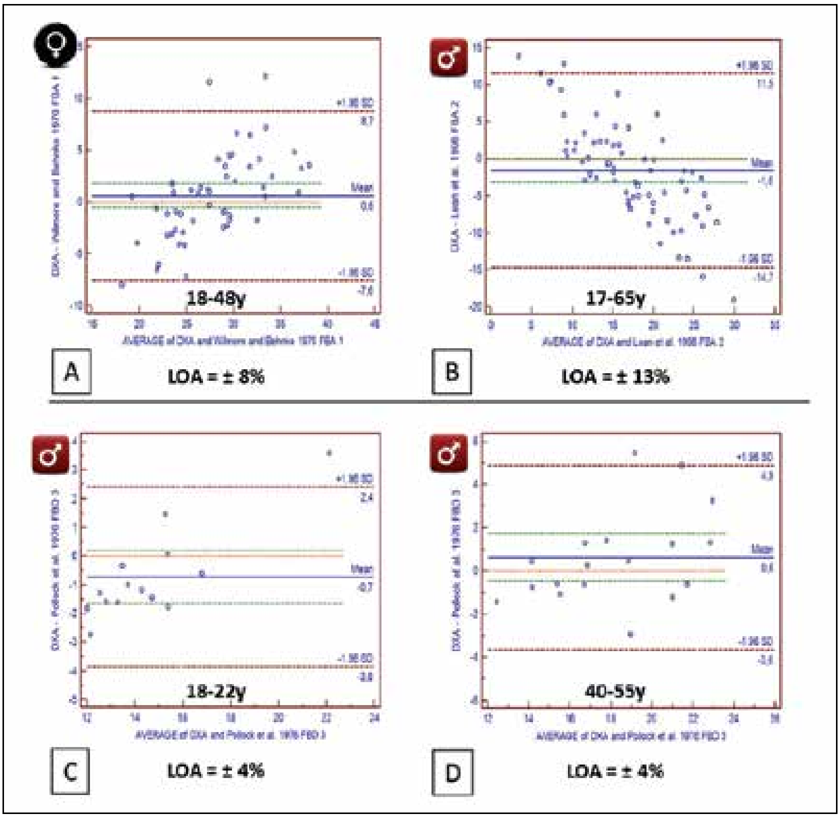
Within this selection of formulae, the Wilmore and Behnke (1970) and Pollock et al (1976; for the age group 40 to 55 years) equations overestimate %AT (Table 4). As concerns the equations of Lean et al. (1996) and Pollock et al (1976; for the age group 18 to 22 years), overestimate %AT (Table 4).

Age range target population	N	♀/♂	Anthropometric equations (reference)	%AT DXA $(\bar{x} \pm SD)$	%AT formulae $(\bar{x} \pm SD)$	Difference DXA-formulae	Pearsons'r
18-48	52	♀	Wilmore et Behnke 1970 FBA	28,17 ± 6,24	27,59 ± 3,90	0,58 ± 4,16	0,76**
18-22	14	♂	Pollock et al. 1976 FBD	14,20 ± 3,29	14,92 ± 1,96	-0,72 ± 1,59	0,93**
40-55	18	♂	Pollock et al. 1976 FBD	18,51 ± 3,78	17,89 ± 2,88	0,62 ± 2,17	0,82**
17-65	69	♂	Lean et al. 1996 FBA	16,25 ± 4,06	17,85 ± 8,89	-1,60 ± 6,68	0,70**

N= Number of subjects; FBA= Anthropometric based formula; FBD= Density based formula; %TA= Percentage adipose tissue; $\bar{x}$= Mean; SD= Standard Deviation; ICC= Intraclass correlation

Table 4. Validity of estimated % body adiposity;

Further analysis with Bland and Altman plots for both ABF and DBF show acceptable to very good mean differences with DXA (from -1.9% up to 1.8%) (Figure 7).



A= Anthropometric based formula of Wilmore & Behnke (1970); B= Anthropometric based formula of Lean et al. (1996); C =Density based formula of Pollock et al. (1976) applied on males aged from 18 to 22 years; D= Density based formula of Pollock et al. (1976) applied on males aged from 40 to 55 years.

Fig. 7. Bland & Altman plots of the selected prediction equations;

The anthropometric based formula of Wilmore and Behnke (1970) (Figure 7A) has a mean/average difference with DXA close to zero ($\bar{x}$ = 0.6) but shows limits of acceptance (between -7.6% and 8.7%). On the graph, we see that this formula tends to overestimate the percentage of body fat compared to DXA for women whose average is less than ± 29% AT and underestimate for women with a body fat percentage greater than 29%.

Conversely, the anthropometric based equation of Lean et al. (1996) (Figure 7B) underestimates the percentage of body fat for men with an average AT% below 17 and tends to overestimate for subjects with an average AT% more than 17.

The density based formula of Pollock et al. (1976) (Figure 7C) for men between 18 and 22 years, defines an average calculated at -0.7 and limits of agreement between -3.8 and 2.4. On the graph, the majority of differences between the two techniques is within the confidence interval for almost all of male subjects included in this formula. Therefore, the results obtained with the formula are relatively close to those obtained with DXA

When this same formula is applied to men with an age between 40 to 55 years (Figure 7D), the limits of agreement are between -3.6 and 4.9 and the average 0.6%AT. The dispersion of values is limited, indicating a close relationship between the results of the equation and those of DXA.

3.3 Discussion

Prediction equations for estimating body fat have been described since 1921 (Matiegka, 1921) and by means of density since 1951 (Brozek and Keys, 1951). Over the years many formulae have been developed in different populations with different characteristics. Today more than 600 prediction equations available to estimate %BF or %AT (e.g. ABF and DBF). The selection of the most appropriate equation for the purpose can be a major concern and must be based on the characteristics of the population on which the chosen equation was validated, realizing that the predictive accuracy of equations remains limited. This indirect approach to BC is automatically exposed to measurement error. In addition the absence of a single gold standard method for obtaining in vivo reference measurements for BC leads to the limited predictive accuracy of field methods in general (Heyward, 1998; Clarys et al., 1999; Clarys et al., 2005).

This study has verified the quality of the predictive accuracy and the applicability of prediction formulae for estimating %AT. Ninety one formulae were retained and applied to a population within different categories of BMI, waist/hip ratio and physical activity level. All formulae were specific to the population at hand and thus a similar outcome of each formula per subject was to be expected. However, such a similarity cannot be confirmed by the results found in this study.

Prediction equations that presented a negative estimation of %AT were excluded. Only 3 out of the original set of > 600 formulae were retained for further analysis with Bland and Altman plots (Figures 7). The plots for all formulae showed an good and acceptable mean/average difference with the DXA result. However the 95% limits of agreement reported in the current study between DXA and the equations are too large. According to Lohman (1986), a 2% accuracy of estimations for %BF is ideal, while a difference in estimated %BF cannot exceed 4.5% (Lohman, 1992). Taking this into consideration, the formulas investigated in this study provide both clinically and biologically unacceptable estimations of %AT. Several authors mentioned previously that prediction equations tend to be inaccurate if the population to which they are applied is different from the one from which the equations were derived (Wilmore and Behnke, 1970; Katch and McArdle, 1973; Sinning, 1978; Vansant et al., 1994; Wong et al., 2000). Brodie et al. (1998) also warns for errors that occur in the calculation of predicted density and its subsequent interpretation as a fat (or correctly termed AT) percentage when using body density as a variable. The well-documented limitations associated with skinfold measurements such as the inability to palpate the fat/muscle interface and the difficulty in obtaining interpretable measurements in obese subjects (Brozek and Kinzey, 1960; Himes et al., 1979; Fanelli and Kuczmarski, 1984) cannot be ignored. Even in a matched population (age, gender, ethnicity, activity level, etc.), the results indicate that formulae are not reliable tools for predicting %BF nor %AT. This lack of observer reliability may be explained on the one hand by the failure of previous studies to report the BMI (underweight, normal weight, overweight, obese), waist/hip ratio distribution categories, as well as by a good definition for athletes (number of hours of sport per week) within a population.

On the other hand, many of the formulae studied, have been initially developed against indirect techniques such as hydrodensitometry, bioelectrical impedance analysis, plethysmography, etc. It can be assumed that errors related to the reference standard used has influenced the difference with DXA.

The use of DXA for assessing BC is not new. Since the seventies total body scans have measured whole-body fat and lean masses in addition to total-body bone mineral content. Lately there has been an increasing usage of DXA as a research tool. This study must make reservation for the DXA comparison. Several studies have shown possible evidence to warn against misinterpretations of DXA data, and it has been suggested that there is insufficient confidence in the ability of DXA to accurately measure the variables it claims to measure (Bolotin, 1998; Bolotin, 2001; Bolotin and Sievanen, 2001; Bolotin et al., 2003; Bolotin, 2004; Bolotin, 2007; Provyn et al., 2008). This lack of confidence has focused principally on the inaccuracies inherent to the surface body density, bone mineral content and the doubtful DXA interpretation of what should be lean body mass (Bolotin, 2001; Bolotin, 2007). Provyn et al. (2008) validated DXA against an in-vitro (dissection) technique and concluded that DXA is a good predictor of %AT, although also warning about the lack of its accuracy when measuring %AT. Because of this critical appraisal, this study has avoided using the wording of “reference standard or golden standard” terminology but has instead referred to a “comparison measure” in the absence of any possible direct measure. Therefore, the choice of an appropriate method and an appropriate prediction equation to precisely assess %AT in individuals remains a challenging task for health and nutrition professionals taking into account that there is no significant difference ($p>0.05$) between the average predicted %AT and the average relative comparison measure.

4. General conclusion

As the result of strong variations in human BC related to age, gender, race, tissue composition and lifestyle, the application of these equations in both men and women, in a clinical setting is questionable (Lohman, 1986; Ortiz et al., 1992; Wong et al., 2000; Clarys et al., 2005). In particular equations resulting from hydrodensitometry-based data collection are subject to some hazards (Scafoglieri et al 2010).

Skinfold compressibility is by no means constant; skin thickness varies with location, females have thinner skin than males and there are significant gender differences in adipose and muscle tissue patterning. Thus, an identical thickness of adipose tissue does not necessarily contain the same concentrations of fat. Despite this variability, a relationship was demonstrated between aggregate skinfold measurements and subcutaneous adipose tissue mass (as opposed to subcutaneous fat) (Clarys et al., 2005).

Prediction equations provide a relatively easy method for the estimation of body adiposity in the context of various health status approaches. The practical "easiness" is due to the use of anthropometry, skinfold thicknesses in particular. Skinfold measurements have been proven adequate (Martin et al., 2003) and thus also the use of a calliper device, which has become a routine laboratory and field instrument that has obtained the status of "tradition" (Martin et al., 1985; Clarys et al., 1987; Clarys et al., 2005). Martin et al. (1985) and Clarys et al. (1987; 2005) showed however that the use of such a skinfold calliper is not without any criticism. Almost all assumptions necessary to convert skinfold calliper readings to percentage ether-extractable fat are clearly unfounded, which means that any whole-body fat prediction that depends on such assumptions cannot be trusted. Having rejected the concept of the prediction of body fat, the next consideration is whether total body adiposity could be confidently predicted from skinfolds instead. To achieve this, SF measurements need to predict subcutaneous adipose tissue adequately, which has been confirmed by a strong relationship between the latter and total body adiposity (Clarys et al., 1999). To pre-empt the question whether equations with anthropometric variables and/or density based on volumes, can predict %BF or %AT, it should repeatedly be made clear that all previously developed formulae produce AT values.

In previously publications, several DBF have shown negative values for %AT in both men and women (Durnin and Womersley 1974; Jackson et al. 1980). In this study, the same phenomenon was observed. These formulae used the Siri equation for predicting %AT and the negative values are a direct consequence of calculated whole-body densities greater than 1.100 g/ml. The occurrence of such values would be a clear indication and confirmation of violation of the assumption of constant density for the FFM. Unrealistic values for %AT (less than 2%), by using the Siri equation, were already reported by Pollock et al. (1977). It can be assumed that many studies encountered negative AT values but probably considered them as erroneous and thus they were never reported.

Subjects were also divided into groups (active and non-active), based on self-report. This method of assessing activity level is questionable due to its subjectivity and the lower ability of adolescents to record their activities. Furthermore, physical activities are generally characterised by irregular bouts of activity of short duration and varied intensity, making it even more difficult to obtain accurate data. In both men and women, prediction equations cannot be used for individual diagnosis but only to give an idea of the BC of an age-, ethnicity-, gender- and activity-matched population.

This study on white females and males confirms that the majority of formulae are not valid for practical use on age-matched individuals. However 5% of the ABF group (e.g. one of the Wilmore and Behnke (1970) formulae and one of the Lean et al. (1996) formulae), has been proven valid on a population corresponding to the original. Of the density derived %AT formulae (DBF), 2% of the developed equations are applicable on similar populations to the original e.g. Pollock et al. (1976 for the age category 18-22y and 40-55y) (Table 4).

Although projects of this kind are time consuming and cumbersome, it is advisable to repeat the same exercise for men and for different ethnic groups for whom whole-body %AT formulae have been developed.

5. References

Bachrach, LK. (2000). "Dual energy X-ray absorptiometry (DEXA) measurements of bone density and body composition: promise and pitfalls." *J Pediatr Endocrinol Metab*,13 Suppl 2, pp. 983-8.

Bolotin, HH. (1998). "Analytic and quantitative exposition of patient-specific systematic inaccuracies inherent in planar DXA-derived in vivo BMD measurements." *Med Phys* 25(2), pp. 139-51.

Bolotin, HH. (2001). "Inaccuracies inherent in dual-energy X-ray absorptiometry in vivo bone mineral densitometry may flaw osteopenic/osteoporotic interpretations and mislead assessment of antiresorptive therapy effectiveness." *Bone* 28(5), pp. 548-55.

Bolotin, HH. (2004). "The significant effects of bone structure on inherent patient-specific DXA in vivo bone mineral density measurement inaccuracies." *Med Phys* 31(4), pp. 774-88.

Bolotin, HH. (2007). "DXA in vivo BMD methodology: an erroneous and misleading research and clinical gauge of bone mineral status, bone fragility, and bone remodelling." *Bone* 41(1), pp. 138-54.

Bolotin, HH. & Sievanen, H. (2001). "Inaccuracies inherent in dual-energy X-ray absorptiometry in vivo bone mineral density can seriously mislead diagnostic/prognostic interpretations of patient-specific bone fragility." *J Bone Miner Res* 16(5), pp. 799-805.

Bolotin, HH.; Sievanen, H. & Grashuis, JL. (2003). "Patient-specific DXA bone mineral density inaccuracies: quantitative effects of nonuniform extraosseous fat distributions." *J Bone Miner Res* 18(6), pp. 1020-7.

Brodie, D.; Moscrip, V.; & Hutcheon, R. (1998). "Body composition measurement: a review of hydrodensitometry, anthropometry, and impedance methods." *Nutrition* 14(3), pp. 296-310.

Brozek, J.; Grande, F.; Anderson, JT. & Keys, A (1963). "Densitometric analysis of body composition: Revision of some quantitative assumptions." *Ann N Y Acad Sci*, 110, pp. 113-40.

Brozek, J. & Keys, A. (1951). "The evaluation of leanness-fatness in man; norms and interrelationships." *Br J Nutr,* 5(2), pp. 194-206.

Brozek, J. & Kinzey, W. (1960). "Age changes in skinfold compressibility." *J Gerontol* 15, pp. 45-51.

Clarys, JP.; Martin, A. & Drinkwater, D. (1988). "Physical and structural distribution of human skin." *Human Biology Budapest,* 18, pp. 55-63.

Clarys, JP.; Martin, AD. & Drinkwater, DT. (1984). "Gross tissue weights in the human body by cadaver dissection." *Hum Biol,* 56(3), pp.: 459-73.

Clarys, JP.; Martin, AD.; Drinkwater, DT. & Marfell-Jones, MJ. (1987). "The skinfold: myth and reality." *J Sports Sci* , 5(1), pp. 3-33.

Clarys, JP.; Martin, AD.; Marfell-Jones, MJ.; Janssens, V.; Caboor, D. & Drinkwater, DT. (1999). "Human body composition: A review of adult dissection data." *Am J Hum Biol,* 11(2), pp. 167-174.

Clarys, JP.; Provyn, S.; Marfell-Jones, M. & Van Roy, P. (2006). "Morphological and constitutional comparison of age-matched in-vivo and post-mortem populations." *Morphologie,* 90(291), pp. 189-96.

Clarys, JP.; Provyn, S. & Marfell-Jones, MJ. (2005). "Cadaver studies and their impact on the understanding of human adiposity." *Ergonomics*, 48(11-14), pp. 1445-61.

Clarys, JP.; Provyn, S.; Wallace, J.; Scafoglieri, A. & Reilly, T (2008). Quality controle of fan beam scanning data processing with in vitro material. Transaction of 2008 IEEE International Conference on industrial engineering and engineering management, Singapore.

Clarys, JP.; Scafoglieri, A.; Provyn, S. & Sesboüé, B. (2009). "The hazards of hydrodensitometry." *Biom Hum et Anthropol*, 27(1-2), pp. 69-78.

Daniel, M.; Martin, AD.; Drinkwater, DT.; Clarys, JP. & Marfell-Jones, MJ. (2003). "Waist-to-hip ratio and adipose tissue distribution: contribution of subcutaneous adiposity." *Am J Hum Biol*, 15(3), pp. 428-32.

Deurenberg, P.; van der Kooy K.; Leenen R.; Weststrate J.A. & Seidell J.C. (1991). Sex and age specific prediction formulas for estimating body composition from bioelectrical impedance: a cross-validation study. *Int J Obes*, 15(1), pp. 17-25.

Duquet, W.; Van Gheluwe, B. & Hebbelinck, M. (1977). "Computer program for calculating Health-Carter anthropometric somatotype." *J Sports Med Phys Fitness,* 17(3), pp. 255-62.

Durnin, JV. & Womersley, J. (1974). "Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged from 16 to 72 years." *Br J Nutr,* 32(1), pp. 77-97.

Edwards, DA. (1951). "Differences in the distribution of subcutaneous fat with sex and maturity." *Clin Sci (Lond)*, 10(3), pp. 305-15.

Evans, EM.; Rowe DA.; Misic MM.; Prior BM. & Arngrimsson SA. (2005). Skinfold prediction equation for athletes developed using a four-component model. *Med Sci Sports Exerc,* 2005. 37(11), pp. 2006-11.

Fanelli, MT. & Kuczmarski, RJ. (1984). "Ultrasound as an approach to assessing body composition." *Am J Clin Nutr,* 39(5), pp. 703-9.

Gallagher, D.; Heymsfield SB.; Heo M.; Jebb SA.; Murgatroyd PR. & Sakamoto Y. (2000). Healthy percentage body fat ranges: an approach for developing guidelines based on body mass index. *Am J Clin Nutr*, 72(3), pp. 694-701.

Garn, SM. (1955). "Relative fat patterning: an individual characteristic." *Hum Biol* 27(2), pp. 75-89.

Garn, SM.; Rosen, NN. & McCann, MB. (1971). "Relative values of different fat folds in a nutritional survey." *Am J Clin Nutr,* 24(12), pp. 1380-1.

Hägar, A. (1991). *Estimation of body fat in infants, children and adolecents. Adipose tissue in childhood.* Bonnet, P. Florida USA, CRC Press Boca Raton, pp. 49-56.

Heath, BH. & Carter, JE. (1967). "A modified somatotype method." *Am J Phys Anthropol,* 27(1), pp. 57-74.

Heyward, VH. (1998). "Practical body composition assessment for children, adults, and older adults." *Int J Sport Nutr,* 8(3), pp. 285-307.

Himes, JH.; Roche, AF. & Siervogel, RM. (1979). "Compressibility of skinfolds and the measurement of subcutaneous fatness." *Am J Clin Nutr,* 32(8), pp. 1734-40.

Jurimae, T.; Jurimae, J.; Wallner, SJ.; Lipp, RW.; Schnedl, WJ.; Moller, R. & Tafeit, E. (2007). "Relationships between body fat measured by DXA and subcutaneous adipose tissue thickness measured by Lipometer in adults." *J Physiol Anthropol,* 26(4), pp. 513-6.

Jurimae, T.; Sudi, K.; Jurimae, J.; Payerl, D.; Moller, R. & Tafeit, E. (2005). "Validity of optical device lipometer and bioelectric impedance analysis for body fat assessment in men and women." *Coll Antropol,* 29(2), pp. 499-502.

Katch, FI. & McArdle, WD. (1973). "Prediction of body density from simple anthropometric measurements in college-age men and women." *Hum Biol,* 45(3), pp. 445-55.

Lean, ME.; Han, TS. & Deurenberg, P. (1996). "Predicting body composition by densitometry from simple anthropometric measurements." *Am J Clin Nutr,* 63(1), pp. 4-14.

Lohman, TG. (1981). "Skinfolds and body density and their relation to body fatness: a review." *Hum Biol*, 53(2), pp. 181-225.

Lohman, TG. (1986). "Applicability of body composition techniques and constants for children and youths." *Exerc Sport Sci Rev,* 14, pp. 325-57.

Lohman, TG. (1992). *Advances in body composition assessment*, Champaign, Ill.: Human Kinetics Publishers.

Madsen, OR.; Jensen, JE. & Sørensen, OH. (1997). "Validation of a dual energy X-ray absorptiometer: measurement of bone mass and soft tissue composition." *Eur J Appl Physiol Occup Physiol,* 75(6), pp. 554-8.

Marfell-Jones, M.; Clarys, JP.; Alewaeters, K. & Martin, AD. (2003). "The hazards of whole body adiposity prediction in men and women." *Biom Hum et Anthropol,* 21(1-2), pp. 103-17.

Martin, AD.; Daniel, M.; Clarys, JP. & Marfell-Jones, MJ. (2003). "Cadaver-assessed validity of anthropometric indicators of adipose tissue distribution." *Int J Obes Relat Metab Disord*, 27(9), pp. 1052-8.

Martin, AD.; Daniel, MZ.; Drinkwater, DT. and Clarys, JP. (1994). "Adipose tissue density, estimated adipose lipid fraction and whole body adiposity in male cadavers." *Int J Obes Relat Metab Disord*, 18(2), pp. 79-83.

Martin, AD.; Janssens, V.; Caboor, D.; Clarys, JP. & Marfell-Jones, MJ. (2003). "Relationships between visceral, trunk and whole-body adipose tissue weights by cadaver dissection." *Ann Hum Biol,* 30(6), pp. 668-77.

Martin, AD.; Ross, WD.; Drinkwater, DT. & Clarys, JP. (1985). "Prediction of body fat by skinfold caliper: assumptions and cadaver evidence." *Int J Obes*, 9 Suppl 1, pp. 31-9.

Martin, R. and Saller, K. (1957). Lehrbuch der anthropologie I, *In systematischer darstellung mit besonderer berücksichtigung der anthropologischen methoden*. Stuttgart, Gustav Fischer Verlag.

Matiegka, J. (1921). "The testing of physical efficiency." *American Journal of Physical Anthropology*, 4, pp. 223-30.

Mueller, WH. (1985). Biology of human fat paterning. London, Ciba foundation.

Mueller, WH. & Stallones, L. (1981). "Anatomical distribution of subcutaneous fat: skinfold site choice and construction of indices." *Hum Biol*, 53(3), pp. 321-35.

Nagy, TR. & Clair, AL. (2000). "Precision and accuracy of dual-energy X-ray absorptiometry for determining in vivo body composition of mice." *Obes Res,* 8(5), pp. 392-8.

Norton, K.; Whittingham, N.; Carter, L.; Kerr, D.; Gore, C. & Marfell-Jones, M. (1996). *International Standards for Anthropometric Assessment*. Underdale, SA, Australia, International Society for the Advancement of Kinanthropometry.

Ortiz, O.; Russell, M.; Daley, TL.; Baumgartner, RN.; Waki, M.; Lichtman, S.; Wang, J.; Pierson, RN. Jr. & Heymsfield, SB. (1992). "Differences in skeletal muscle and bone mineral mass between black and white females and their relevance to estimates of body composition." *Am J Clin Nutr,* 55(1), pp. 8-13.

Pollock, ML.; Gettman, LR.; Jackson, A.; Ayres, J.; Ward, A. & Linnerud, AC. (1976). "Body composition of elite class distance runners." *Ann N Y Acad Sci,* 301, pp. 361-70.

Provyn, S.; Clarys, JP.; Wallace, J.; Scafoglieri, A. and Reilly, T. (2008). "Quality control, accuracy, and prediction capacity of dual energy X-ray absorptiometry variables and data acquisition." *J Physiol Anthropol,* 27(6), pp. 317-23.

Provyn, S.; Scafoglieri A.; Tresignie J.; Bautmans I. ; Reilly T. & Clarys JP. (2011) Quality control of 157 whole body adiposity prediction formulae in age and activity matched men. *J Sports Med Phys Fitness,* Sep;51(3), pp.426-34.

Provyn, S.; Wallace J. ; Scafoglieri A. ; Sesboüé B. ; Marfell-Jones M. ; Bautmans I. & Clarys JP. (2010), Formules de prédiction de l'adiposité chez la femme – contrôle de qualité Original Research Article, *Science & Sports*, Volume 25, Issue 6, December , pp.291-303

Rolland-Cachera, MF. & Brambilla, P. (2005). "Reference body composition and anthropometry." *Int J Obes (Lond),* 29(8), pp. 1010.

Scafoglieri A.; Provyn S.; Bautmans I.; Wallace J.; Sutton L.; Tresignie J.; Louis O.; De Mey J. & Clarys JP. (2010). *Critical Appraisal of Data Acquisition in Body Composition: Evaluation of Methods, Techniques and Technologies on the Anatomical Tissue-System Level,* Data Acquisition, Michele Vadursi (Ed.), ISBN: 978-953-307-193-0, InTech,

Sinning, WE. (1978). "Anthropometric estimation of body density, fat, and lean body weight in women gymnasts." *Med Sci Sports,* 10(4), pp. 243-9.

Siri, WE. (1956). "The gross composition of the body." *Adv Biol Med Phys,* 4, pp. 239-80.

Snijder, MB.; Visser, M.; Dekker, JM.; Seidell, JC.; Fuerst, T.; Tylavsky, F.; Cauley, J.; Lang, T.; Nevitt, M. & Harris, TB. (2002). "The prediction of visceral fat by dual-energy X-ray absorptiometry in the elderly: a comparison with computed tomography and anthropometry." *Int J Obes Relat Metab Disord,* 26(7), pp. 984-93.

Tafeit, E.; Moller, R.; Jurimae, T.; Sudi, K. & Wallner, SJ. (2007). "Subcutaneous adipose tissue topography (SAT-Top) development in children and young adults." *Coll Antropol*, 31(2), pp. 395-402.

Vansant, G.; Van Gaal, L. & De Leeuw, I. (1994). "Assessment of body composition by skinfold anthropometry and bioelectrical impedance technique: a comparative study." *JPEN J Parenter Enteral Nutr*, 18(5), pp. 427-9.

Verbraecken, J.; Van de Heyning, P.; De Backer, W. & Van Gaal, L. (2006). "Body surface area in normal-weight, overweight, and obese adults. A comparison study." *Metabolism,* 55(4), pp. 515-24.

Wallace, J.; George, K. and Reilly, T. (2005). "Validation of dual-energy x-ray absorptiometry for segmental body composition analysis." *Journal of Sports Sciences,* 11/12(23), pp. 1165-6.

Wilmore, JH. & Behnke, AR. (1970). "An anthropometric estimation of body density and lean body weight in young women." *Am J Clin Nutr* 23(3), pp. 267-74.

Wong, WW.; Stuff, JE.; Butte, NF.; Smith, EO. & Ellis, KJ. (2000). "Estimating body fat in African American and white adolescent girls: a comparison of skinfold-thickness equations with a 4-compartment criterion model." *Am J Clin Nutr,* 72(2), pp. 348-54.

8

Physical Activity, Inactivity, and Nutrition Behavior Among Children: Investigating Compensation and Transfer Effects

Judith Väth[1], Katie Amato[2] and Claudio R. Nigg[2]
[1]Institute for Sport and Sport Sciences, University of Karlsruhe, Karlsruhe
[2]Department of Public Health Sciences, University of Hawai'i at Mānoa, Honolulu, HI
[1]Germany
[2]USA

1. Introduction

The increasing prevalence of overweight children is an important public health problem in the United States. Nearly 1/3 of children are considered to be at risk for being overweight and currently more than 9 million children over 6 years of age are considered obese (Ogden et al. 2002; Koplan et al. 2005). Obesity and being overweight is a risk factor for several diseases: type 2 diabetes, cardiovascular disease, hypertension, osteoporosis, and certain types of cancer (Eaton et al. 2006).

Both physical activity (PA) and nutrition behaviors have been shown to be an important and effective method to reduce weight. Physical activity expends energy and can lead to a reduction in weight loss. A meta-analysis showed that there is a small to moderate relationship between body fat and activity in children (Rowlands, Ingledew and Eston 2000). But to reduce adiposity in children and adolescents of normal weight an intense level of PA for a longer duration is needed (Barbeau and Litaker 2003; Eliakim et al. 2000). Time spent in vigorous and hard activity correlated significantly with percentage of body fat but not with BMI in 5-10.5 year-olds (Abbott and Davies 2004). Physical activity not only positively influences physiological factors, but also has a positive effect on psychological aspects. Regular PA can increase the ability to cope with stress and lead to an improved health perception and quality of life (Röthlosberger, Calmonte and Seiler 1997). Strong et al. (2005) emphasized many other beneficial effects of PA, such as better cardiovascular health and self-esteem.

Most research examining physical inactivity focuses on television (TV) viewing. Some cross-sectional studies found positive associations between TV viewing and obesity. An analysis of the CDC 1999 Youth Risk Behavior Survey demonstrated a significant association between overweight and viewing TV more than 2 hours per day (Lowry et al. 2002). Also, eating meals in front of the TV may influence energy intake because it is associated with lower fruit, vegetable, and juice intake and greater intake of salty snacks, pizza, soft drinks, and red meat (Proctor 2003; Coon et al. 2001).

Some clinical evidence shows that receiving advice to increase fruit and vegetable consumption is an effective strategy for weight management as fruits and vegetables have a low energy density, are high in fiber, and may cause satiety. In addition, consumption of fruits and vegetables could also displace consumption of less healthy and higher energy-dense foods (Sherry 2005). Children and adolescents in the US have not consumed the recommended 5 servings of fruits and vegetables per day (Cavadini, Siega-Rizz and Popkin 1996). Studies report that only 18% of girls and 14% of boys consume the recommended number of servings of fruits and vegetables (American Dietetic Association 2004; Enns, Mickle and Goldmann 2003).

Relationships between PA, inactivity, and nutrition behavior are consistently shown in studies on elementary school children (Sallis, Prochaska and Taylor 2000; Driskell et al. 2008; Pearson and Biddle 2011). For example, in a comprehensive review of PA correlates among children Driskell et al. (2008) found that healthy diet, intention to be active, and PA preferences (among others) cluster with PA. Traditionally, PA and nutrition interventions have focused on influencing single behaviors. However, recent research suggests that multiple behavior change interventions may have a greater impact than single behavior change interventions (Nigg, Allegrante and Ory 2002; Emmons et al. 1994). Because these behaviors are associated in individuals; combined PA, inactivity, and nutrition interventions hold promise for effectively influencing multiple outcomes. In a PA and healthy eating intervention evaluation on adults, a multiple behavior intervention was 3 times as successful as a single behavior intervention (Johnson et al. 2008). In another study on children, a combined nutrition and PA group scored significantly better than a control group on measures of nutrition knowledge. Results of single- and multiple behavior change interventions imply that future investigation of how changes in PA, inactivity, and nutrition may impact each other is warranted (Warren et al. 2003).

Physical activity, inactivity, and nutrition behaviors may act as gateway behaviors. Gateway behaviors are those which when changed lead to a positive change in another health behavior (Nigg et al. 2009). Depending on the interaction of the behaviors, change in more than one behavior may be due to transfer or compensation effects. Borrowed from learning and teaching research (Barnett and Ceci 2002), transfer effects describe the translation of knowledge and confidence in one health behavior to another. Research on transfer effects is inconclusive: some studies report null results (Ussher, Taylor and Faulkner 2008; Wilcox et al. 2000), while others provide support (Nigg et al. 2009; Fleig et al. 2011). Transfer effects may depend on the 1) co-occurrence of behaviors 2) similarity of health behavior domains (Flay and Petraitis 1994), and 3) individual ability to transfer skills to another domain (Perkins and Saloman 1994).

Because PA, inactivity, and nutrition are associated and naturally co-occur, transfer effects may explain the success of behavior change interventions. To date, transfer effects in these domains have not yet been investigated in children. Research on single behavior interventions measuring other behaviors has found support for transfer effects between PA and nutrition in adults. One study found that self-efficacy in exercise served as a gateway to healthy diet (Tucker and Reicks 2002). Additionally, a recent exercise intervention found PA transfer effects on fruit and vegetable intake (Fleig et al. 2011). Combined PA and nutrition interventions evaluating transfer effects are few. One intervention comparing single versus multiple behavior interventions that targeted PA and fat intake found that success at improving both behaviors was not associated with intervention condition (Vandelanotte et al. 2008). Results suggest that participants who successfully changed both behaviors in the single behavior change intervention experienced transfer effects.

In contrast to transfer effects, the opposite interaction effect may occur where individuals may compensate for their risk behavior by performing another health behavior (see Compensatory Health Belief Model, Knauper et al. 2004). For example, it has been shown that smokers are more physically active in order to compensate for their unhealthy lifestyle (Xu 2002). Less literature is available on compensation effects of PA, inactivity, and nutrition. One study exploring the effects of a PA intervention on adult's nutrition found that PA was not a gateway behavior for fruit and vegetable consumption. Instead, they found that increases in PA activity were associated with increases in fat intake (Dutton et al. 2008) suggesting that participants may have compensated for increased fat intake with increased PA. Both transfer and compensation effects of PA, inactivity, and nutrition behavior may have important intervention implications, but they have not yet been explored in children. Therefore, the purpose of the current study is to compare PA, inactivity, and nutrition behavior, their influence on each other in children, and to determine if there are compensation or transfer effects.

2. Methods

Twenty-one schools in the state of Hawaii (the islands of Oahu, Hawaii, and Maui) participating in Fun 5 were randomly selected and stratified by afterschool care provider, year joined, and county. Fun 5 is a nutrition and PA program aimed at reducing obesity through increasing fruit and vegetable consumption and PA (Battista et al. 2005). Parental consent and student surveys containing PA, inactivity, and nutrition behavior questions were sent to the site coordinators with instructions for distribution, administration, and return of completed materials. One site did not return any surveys and 7 sites did not get parental permission. A significant portion of sites' student surveys (N= 250) were obtained without consent and destroyed. In the final analysis, only students from 13 sites were evaluated. There were two measurement points: one at the beginning of the school year (Fall 2005 – baseline T1); and the other at the end of the school year (Spring 2006 – follow-up T2). The University of Hawaii Committee on Human Subjects approved this research.

2.1 Participants

Participants were enrolled in Hawaii's A-Plus public elementary after-school program (A+). A+ is a state mandated after-school program for children in public elementary schools that begins immediately after the end of the school day until the last child is picked-up (2pm~5:30pm). The program includes snack time, homework time, PA, and enrichment activities such as arts and crafts.

At T1, 188 student surveys from 13 sites were available for analysis (53.2% female; grade 4= 39%; grade 5= 34%; and grade 6= 27%). At T2, 137 students (56.7% female; grade 4= 37%; grade 5= 37%; and grade 6= 26%) completed the survey from the same 13 sites. For the analysis, only children who took part on both measurement points are included. Table 1 shows the sample of those participants that completed T2 and those that did not complete T2.

	Completed (N= 137)		Not completed (N= 51)		t-test	
Variable	mean	Std.	mean	Std.	df	sig.
Strenuous PA (min/wk)	202.28	135.86	213.14	137.18	185	.628
Moderate PA (min/wk)	147.45	125.47	152.24	137.72	184	.823
Mild PA (min/wk)	119.48	140.75	136.67	141.08	181	.469
Inactivity (hrs/day)	4.47	3.252	3.88	3.090	186	.268
Fruit (serv/day)	4.69	2.950	4.32	2.622	180	.438
Vegetable (serv/day)	3.77	2.825	3.82	2.561	179	.915

Note: PA – physical activity; min – minutes; wk – week; hrs – hours; serv – servings
Std. – Standard deviation; df – difference; sig – significance; N – numbers.

Table 1. Mean, standard deviation and significance of PA, inactivity, and nutrition variables at T1 (completed or not completed T2)

2.2 Measures

The student survey included measures on demographics (grade and gender), PA, inactivity, and fruit and vegetable consumption. An adaptation of Godin & Shephard's Leisure-Time Physical Activity Questionnaire (Godin and Shepard 1985) indicates how many days during an average week people are engaged in strenuous, moderate, and mild PA during their free time. Strenuous PA is defined as "heart beats rapidly, sweating" examples are: running, jogging, soccer, squash, cross country skiing, judo, roller skating, vigorous swimming, vigorous long distance bicycling, vigorous aerobic dance classes, and/or heavy weight training. Moderate PA is defined as "not exhausting, light sweating" examples are: fast walking, baseball, tennis, easy bicycling, volleyball, badminton, easy swimming, popular, folk and / or hula dancing. Mild PA is defined as "minimal effort, no sweating, e.g. easy walking, yoga, archery, fishing, bowling, lawn bowling, shuffleboard, horseshoes, and/ or golf. For a sample of adults (Godin and Shepard 1985) aged 18-65 years, two-week test-retest reliabilities of .94, .46, and .48 were reported for strenuous, moderate, and mild PA respectively. Strenuous PA was significantly associated with maximum oxygen intake (VO2max; r = .38) and percentage of body fat (r = .21). The instrument was found to be significantly related to caloric accelerometer readings (r = .32), metabolic equivalents (METs; r = .36), treadmill PA time (r = .57), percentage of body fat (r = -.43), VO2max (r = .56), and the stages of PA across populations (Jacobs et al. 1993; Lee et al. 2001; Schumann et al. 2002). A second set of questions reflects the minutes (in 10 min increments) that participants spent in each activity level each day. This allows a calculation of min/week of PA for each intensity level.

For sedentary behavior one question addressed how many hours the student watched TV or played video games on an average day (Buckworth and Nigg 2004). Validity of this item has shown a small negative correlation in PA with children (Nigg 2005).

Participants reported how many servings of fruits and how many servings of vegetables they ate each day. A serving was described as: ½ cup of cooked vegetables = size of 2 ping-pong balls; 1 cup of salad = size of 1 baseball; 1 piece of fruit = size of 1 baseball; or ¾ cup of 100% fruit juice = 6 ounces. The single items addressing the average number of fruits and the average number of vegetables eaten each day have documented validity and reliability in adolescents (Prochaska 2000).

2.3 Analysis

Data analysis was conducted via SPSS 14.0 (2005, SPSS Worldwide Headquarters, Chicago, IL). Three variables represent PA level in min/week (strenuous, moderate, mild), one variable measures the inactivity (TV watching or videogame playing), and two variables address healthy nutrition behavior (fruit and vegetable consumption). For the measurement of strength and direction of cross-sectional relationships Pearson Correlations were conducted. Cohen's guidelines (1988) for interpretation of the correlation coefficient were used: small ranging from │.10│ - │.29│; medium ranging from │.30│ - │.49│; and large ranging from │.50│ – │1.0│. A linear regression was conducted to assess the longitudinal relationship between a dependent variable, independent variables, and a random term. Independent variables were all the behaviors at T1. For example, strenuous PA, moderate PA, mild PA, inactivity, fruit and vegetables at T1 were used to predict one dependent variable at T2 (e.g., strenuous PA).

3. Results

A t-test was conducted to compare those who completed T2 with those who did not complete T2. Means and standard deviations are represented in Table 1 and show for all six variables that there is no significant difference between both groups ($p > .05$). Only the 137 participants who completed T1 and T2 were included in the analysis. Cross-sectional analysis with all the participants at T1 did not alter any conclusions (results not shown). Table 2 shows the mean and standard deviation of all variables across time. There were no significant differences between T1 and T2 ($p > .05$).

	T1 (Fall 2005)		T2 (Spring 2006)		t-test	
Variable	mean	std.	mean	Std.	df	Sig.
Strenuous PA (min/wk)	202.3	135.9	214.2	134.8	133	.430
Moderate PA (min/wk)	147.4	125.5	147.1	127.0	135	.971
Mild PA (min/wk)	119.5	140.7	126.1	142.0	132	.530
Inactivity (hrs/day)	4.47	3.252	4.26	3.094	135	.606
Fruit (serv/day)	4.69	2.950	4.18	2.784	125	.111
Vegetable (serv/day)	3.77	2.825	3.77	2.577	124	.842

Note: PA – physical activity; min – minutes; wk – week; hrs – hours; serv – servings Std. – Standard deviation; df – difference; sig – significance

Table 2. Mean, standard deviation and significance of PA, inactivity, and nutrition variables

The correlations between the variables of PA, inactivity, and nutrition behavior of T1 are represented in Table 3 and Table 4 shows the same correlations for T2. Regarding T1: there is a small correlation between strenuous PA and fruit consumption (r= .256**), between mild

PA and inactivity (r= .181*), between mild PA and fruit consumption (r= .229*) and also between mild PA and vegetable consumption (r= .248*). A medium correlation could be found between strenuous and mild PA (r= .366**), between moderate and mild PA (r= .419**), between moderate PA and fruit consumption (r= .370**), between vegetable consumption and strenuous PA (r= .337**) and moderate PA (r= .379**). Large relationships between strenuous and moderate PA (r= .558**) and between fruit and vegetable consumption (r= .624**) were found.

Pearson's correlation	Strenuous PA (min/wk T1)	Moderate PA (min/wk T1)	Mild PA (min/wk T1)	Inactivity (hrs/day T1)	Fruit (serv/day T1)	Vegetable (serv/day T1)
Strenuous PA (min/wk T1)		.558**	.366**	-.003	.256**	.337**
Moderate PA (min/wk T1)			.419**	.084	.370**	.379**
Mild PA (min/wk T1)				.181*	.229*	.248*
Inactivity (hrs/day T1)					.069	.008
Fruit (serv/day T1)						.624**
Vegetable (serv/day T1)						

**. Correlation is significant at the 0.01 level (2-tailed)
*. Correlation is significant at the 0.05 level (2-tailed)
Note: PA – physical activity, min – minutes, wk – week, hrs – hours, serv – servings

Table 3. Correlation between PA, inactivity, and nutrition variables of T1

Pearson's Correlation	Strenuous PA (min/wk T2)	Moderate PA (min/wk T2)	Mild PA (min/wk T2)	Inactivity (hrs/day T2)	Fruit (serv/day T2)	Vegetable (serv/day T2)
Strenuous PA (min/wk T2)		.304**	.274**	-.081	.288**	.163
Moderate PA (min/wk T2)	.		.520**	-.044	.020	.006
Mild PA (min/wk T2)				.043	.052	.054
Inactivity (hrs/day T2)					.062	-.129
Fruit (serv/day T2)						.538**
Vegetable (serv/day T2)						

**. Correlation is significant at the 0.01 level (2-tailed)
*. Correlation is significant at the 0.05 level (2-tailed)
Note: PA – physical activity, min – minutes, wk – week, hrs – hours, serv - servings

Table 4. Correlation between PA, inactivity, and nutrition variables of T2

Regarding T2: there is a small relationship between mild and strenuous PA (r= .274**), between strenuous PA and fruit consumption (r= .288**). A medium correlation is evidenced for strenuous and moderate PA (r= .304**). A large relationship was found between moderate and mild PA (r= .520**) and between fruit and vegetable consumption (r= .538**).

Dependent Independent	Strenuous PA (min/wk T2) (r^2=.104)	Moderate PA (min/wk T2) (r^2= .174)	Mild PA (min/wk T2) (r^2= .112)	Inactivity (hrs/day T2) (r^2= .156)	Fruit (serv/day T2) (r^2= .263)	Vegetable (serv/day T2) (r^2= .180)
Strenuous PA	.186	.298	-.015	.000	.000	.003
(min/wk T1)	*.081*	*.002*	*.894*	*.921*	*.823*	*.179*
Moderate PA	.226	.017	.238	-.004	-.004	-.004
(min/wk T1)	*.062*	*.879*	*.062*	*.108*	*.107*	*.100*
Mild PA	-.055	.168	.189	.002	.001	.000
(min/wk T1)	*.569*	*.057*	*.070*	*.393*	*.629*	*.953*
Inactivity	-1.035	-3.884	1.765	.312	.050	.083
(hrs/day T1)	*.777*	*.250*	*.655*	*.000*	*.483*	*.244*
Fruit	3.371	2.794	1.394	.168	.611	.056
(serv/day T1)	*.505*	*.543*	*.794*	*.139*	*.000*	*.576*
Vegetable	-4.574	-11.346	-.976	-.182	-.235	.336
(serv/day T1)	*.403*	*.022*	*.865*	*.133*	*.027*	*.002*

Note: PA – physical activity, min – minutes, wk – week, hrs – hours, serv - servings

Table 5. Linear Regression of PA, inactivity, and nutrition variables (unstandardized Coefficient B, *significance*)

The outcomes of the linear regression analyses (see Table 5) show that there is no significant ($p > .05$) relationship of T2 strenuous PA and of T2 mild PA with all the T1 predictors. However, T2 strenuous PA is marginally ($p \leq .10$) significant with T1 strenuous PA and T1 moderate PA. T2 mild PA is also marginally ($p \leq .10$) significant with T1 moderate PA and T1 mild PA. There is a significant ($p \leq .05$) relationship between T2 moderate PA and T1 strenuous PA (unstandardized coefficient B= .298) and marginally with T1 mild PA (unstandardized coefficient B= .168). Also, a significant negative relationship was found between T2 moderate PA and T1 vegetables (unstandardized coefficient B= -11.346). For T2 fruit consumption there is a significant negative relationship with T1 vegetable consumption (unstandardized Coefficient B= -.235). T2 vegetable consumption and inactivity are only significant predictors of themselves at T1.

4. Discussion

The purpose of the current study was to compare PA, inactivity, nutrition behaviors, and their influence on each other in children to determine if there are compensation or transfer effects between the different domains of behavior. According to the Compensatory Health Belief Model, individuals can compensate their risk behavior by performing another health behavior (Knauper et al. 2004).The opposite effect occurs when individuals transfer their knowledge and experiences from one behavior change to another (Barnett and Ceci 2002).

The results of the cross-sectional analyses show that PA behavior relates to itself and it relates to fruit and vegetable consumption. Students that are strenuously active are more likely to be moderately active. This may be because they warm-up with a moderate intensity activity. Students that are mildly active are also moderately active but are less strenuously active. Regarding the relationship between PA and nutrition behavior, the outcomes of T1 show a strong relationship between moderate PA and fruit and vegetable consumption. Also, strenuous PA is related to fruit and vegetable consumption. The results support a transfer effect because very active children are also more likely to eat more fruits and vegetables.

Regarding the relationship between PA and fruit and vegetable consumption of T2, there is only a small correlation between strenuous PA and fruit consumption. This result also points to a transfer effect. The longitudinal results are weaker: they confirm the transfer effect within PA intensities but not for the transfer to fruit and vegetable consumption.

Cross-sectional and longitudinal analysis shows an independence of the inactivity variable both from PA and from the nutrition variables. Similarly, Anderson et al. (1998) and Nigg et al. (2002) did not find any meaningful relationship between TV watching and PA. Olivares et al. found no association between nutritional status and television viewing (Olivares et al. 2004). However, studies have shown that TV viewing may contribute to a decline in fruit and vegetable consumption among adolescents. Boynton-Jarrett et al. (2003) documented that the fruit and vegetable consumption was negatively associated with hours of TV viewing. Their prospective analyses indicated that both baseline television viewing and change in television viewing independently predicted a reduction in fruit and vegetable consumption.

Cross-sectional and longitudinal results of fruit and vegetable consumption show an expected transfer effect within fruits and vegetables as both variables represent eating behaviors. There is also cross-sectional indication of transfer effects between fruit and vegetable consumption and PA; however, this was not confirmed by the longitudinal analysis. There were no predictive relationships with PA or inactivity. As expected, the relationship is weaker over time as people change their behavior. The longitudinal results also show a negative prediction of fruits by vegetables. This negative relationship between the two nutrition variables has statistical reasons and points to a suppression effect. A statistical suppression means that instead of the drop that we would see from the direct effect of the same behavior on the outcome when vegetable variable is introduced, the opposite happens. The testing of suppression should be evidenced on a priori assumptions about the theoretical relation between the variables and the role of the second predictor variable as a suppressor.

4.1 Limitations

There are some limitations that need to be considered when interpreting the outcomes of this study. First, data was collected by self-report only and thus, the outcomes could possibly be influenced by social desirability. Second, the survey only included questions concerning PA of 10 minutes or more of continuous activity. Corbin and Pangrazi (1998) noted that children normally exercise in short bouts of activity rather than being constantly active. Third, the sample size (N= 137) did not allow for subgroup analysis (e.g., gender). With decreasing numbers the chance of random distortion of the results increases (Boes, Haensel and Schott 2000). Lastly, the large amount of students who did not provide data influenced the representativeness. However, as we were interested in variable relationship and not prevalence we do not deem this a serious limitation.

5. Conclusion

In conclusion, there seems to be no compensation effect between PA, inactivity, and nutrition behavior in children. Inactivity does not seem to be related to PA or nutrition. However, there are important transfer effects between different PA intensities and between nutrition behaviors in children. These outcomes support growing evidence that multiple behavior interventions have the potential for a greater impact on public health than single behavior interventions. More research is needed to understand the mechanisms and moderations (e.g., gender) of the influence and relationship between PA, inactivity, and nutrition behavior. Finally, intervention research should investigate how to capitalize and promote the important transfer effects.

6. Acknowledgements

This research was funded by the Hawaii Medical Service Association, an independent licensee of the Blue Cross and Blue Shield Association. We would like to thank the valuable contributions of: Marisa Yamashita, Jo Ann Chang, Richard Chung, Cynthia Ross, Paula Adams, Phoebe Hwang, Michele Hamada, and Jessica Westling; the private providers (Kama'aina Kids, YMCA, and the Hawaii State Department of Education) and the participating sites.

7. References

Abbott, R.A. and Davies, P.S.W. (2004). "Habitual Physical Activity and Physical Activity Intensity: Their Relation to Body Composition in 5.0-10.5 Year Old Children." *European Journal of Clinical Nutrition* 58: 285-291.

American Dietetic Association. (2004). "Position of the American Dietetic Association: Dietary Guidance for Healthy Children Aged 2 to 11 years." *Journal of the American Dietetic Association* 104: 66-667.

Anderson, R.E., Crespo, C., Bartlett, S.J., Cherskin, L.J. and Pratt, M. (1998). "Relationship of Physical Activity and Television Watching with Body Weight and Level of Fatness Among Children: Results from the Third National Health and Nutrition Examination Survey." *JAMA* 279: 938-942.

Barbeau, P. and Litaker, M.S. (2003) "Effect of a Physical Activity Intervention on Bone Mineral Density in Black Girls: Relative Contribution of Changes in Gar Mass and Fat Free Soft tissue." *Medicine & Science in Sports & Exercise* 35: S82.

Barnett, S.M. and Ceci, S. J. (2002). "When and Where Do We Apply What We Learn?: A Taxonomy for Far Transfer." *Psychological Bulletin* 128(4): 612-637

Battista, J., Nigg, C.R., Chang, J.A., Yamashita, M. and Chung, R. (2005). "Elementary After School Programs: An Opportunity to Promote Physical Activity for Children." *California Journal of Health Promotion* 3: 108-180.

Boes, K., Haensel, F. and Schott, N. (2000). "Empirical Investigations in the Sport Science." Hamburg, Germany: Czwalina.

Boynton-Jarrett, S.M., Thomas, T.N., Peterson, K.E., Wiecha, J., Sobol, A.M. and Gortmarker, S.L. (2003) "Impact of Television Viewing Patterns on Fruit and Vegetable Consumption among Adolescents." *Pediatrics* 112: 1321-1326.

Buckworth, J. and Nigg, C.R. (2004). "Physical Activity, Exercise, and Sedentary Behavior in College Students." *Journal of American College Health* 53: 28-34.

Cavadini, C., Siega-Riz, A.M. and Popkin, B.M. (2000). "U.S. Adolescent Food Intake Trends from 1965 to 1996." *Archives of Disease in Children* 83: 18-24.

Cohen, J. (1988). "Statistical Power Analysis for the Behavioral Sciences." 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Coon, K.A., Goldberg, J., Rogers, B.L. and Tucker, K.L. (2001). "Relationships Between Use of Television During Meals and Children's Food Consumption Patterns." *Pediatrics* 107: E7.

Corbin, C.B. and Pangrazi, R.P. (1998) "Physical Activity for Children: A Statement of Guidelines." National Association for Sport and Physical Education Publications. Reston: VA.

Driskell, M.M., Dyment, S., Mauriello, L., Castle, P. and Sherman, K. (2008). "Relationships Among Multiple Behaviors for Childhood and Adolescent Obesity Prevention." *Preventive Medicine* 46(3): 209-215.

Dutton, G.R., Napolitano, M.A., Whiteley, J.A. and Marcus, B.H. (2008). "Is Physical Activity a Gateway Behavior for Diet? Findings from a Physical Activity Trial." *Preventive Medicine* 46(3): 216-221.

Eaton, D.K., Kann, L., Kinchen, S., Ross, J., Hawkins, J., Harris, W.A., Lowry, R., McManus, T., Chyen, D., Shanklin, S., Lim, C., Grunbaum, J.A. and Wechsler, H. (2006). "Youth Risk Behavior Surveillance—United States, 2005." *Journal of School Health* 76(7): 353-372.

Eliakim, A., Makowski, G.S., Brasel, J.A. and Cooper D.M. (2000) "Adiposity, Lopeid Levels, and Brief Endurance Training in Non-obese Adolescent Males." *International Journal of Sports Medicine* 21: 332-337.

Emmons, K.M., Marcus, B.H., Linnan, L., Rossi, J.S. and Abrams, D.B. (1994). "Mechanisms in Multiple Risk Factor Interventions: Smoking, Physical Activity, and Dietary Fat Intake Among Manufacturing Workers. Working Well Research Group." *Preventive Medicine* 23(4): 481-489.

Enns, C.W., Mickle, S.J. and Goldmann, J.D. (2003). "Trends in Food and Nutrient Intakes by Adolescents in the United States." *Family Economics and Nutrition Review* 15: 15-27.

Flay, B.R. and Petraitis, J. (1994). "The Theory of Triadic Influence: A New Theory of Health Behavior with Implications for Preventive Interventions." *Advances in Medical Sociology, Vol. IV, A Reconsideration of Models of Health Behavior Change*. In: G. Albrecht (ed). Greenwich, CT: JAI Press: 19–44

Fleig, L., Lippke, S., Pomp, S. and Schwarzer, R. (2011). "Intervention Effects of Exercise Self-regulation on Physical Exercise and Eating Fruits and Vegetables: A Longitudinal Study in Orthopedic and Cardiac Rehabilitation." *Preventive Medicine* 53(3): 182-187.

Godin, G. and Shepard, R.J. (1985). "A Simple Method to Assess Exercise Behavior in the Community." *Canadian Journal of Applied Sport Sciences* 10: 141-146.

Jacobs, D.R., Ainsworth, B.E., Hartman, T.J. and Leon, A.S. (1993). "A Simultaneous Evaluation of 10 Commonly Used Physical Activity Questionnaires." *Medicine & Science in Sports & Exercise* 25: 81-91.

Johnson, S.S., Paiva, A.L., Cummins, C.O., Johnson, J.L., Dyment, S.J., Wright, J.A. and Sherman, K. (2008). "Transtheoretical Model-based Multiple Behavior Intervention for Weight Management: Effectiveness on a Population Basis." *Preventive Medicine* 46(3): 238-246.

Knauper, B., Rabiau, M., Cohen, O. and Patriciu, N. (2004). "Compensatory Health Beliefs: Scale Development and Psychometric Properties." *Health Psychology* 19: 607-624.

Koplan, J.P., Liverman, C.T. and Kraak, V.I. (2005). "Preventing Childhood Obesity: Health in the Balance." Washington, DC: National Academies Press.

Kuczmarski, R.J., Ogden, C.L., Guo, S.S., Grummer-Strawn, L.M., Flegal, K.M., Mei, Z., Curtin, L.R., Roche, A.F. and Johnson, C.L. (2002). "2000 CDC Growth Charts for the United States: Methods and Development." *Vital Health Stat 11* 246: 1-190.

Lee, R.E., Nigg, C.R., DiClemente, C.C. and Courneya, K.S. (2001). "Validating Motivational Readiness for Exercise Behavior with Adolescents." *Research Quarterly for Exercise and Sport* 72: 401-410.

Lowry, R., Wechsler, H., Galuska, D.A., Fulton, J.E. and Kann, L. (2002). Television Viewing and its Associations with Overweight, Sedentary Lifestyle, and Insufficient Consumption of Fruits and Vegetables Among U.S. High School Students: Differences by Race, Ethnicity, and Gender." *Journal of School Health* 72: 413-421.

Nigg, C.R. (2005). "There Is More to Stages of Exercise Than Just Exercise." *Exercise and Sport Science Reviews* 33: 32-35.

Nigg, C.R., Lee, H. R., Hubbard, A.E. and Min-Sun, K. (2009). "Gateway Health Behaviors in College Students: Investigating Transfer and Compensation effects." *Journal of American College Health* 58(1): 39-44.

Nigg, C.R., Allegrante, J.P. and Ory, M. (2002). "Theory Comparison and Multiple-behavior Research: Common Themes Advancing Health Behavior Research." *Health Education Research* 5: 670-679

Ogden, C.L., Flegal, K.M., Carroll, M.D. and Johnson, C.L. (2002). "Prevalence and Trends in Overweight Among U.S. Children and Adolescents: 1999, 2000." *JAMA* 144: 1728-1732.

Olivares, S., Kain, J., Lera, L., Pizarro, F., Vio, F. and Moron, C. (2004). "Nutritional Status, Food Consumption and Physical Activity Among Chilean School Children: a Descriptive Study." *European Journal of Clinical Nutrition* 58: 1278-1285.

Pearson, N. and Biddle, S.J.H. (2011). "Sedentary Behavior and Dietary Intake in Children, Adolescents, and Adults: A Systematic Review." *American Journal of Preventive Medicine,* 41(2): 178-188.

Perkins, D. and Saloman, G. (1994) "Transfer of Learning." T. Husen and T. Postlethwaite (eds). *International Encyclopedia of Education Second Edition* Oxford: Pergamom Press

Prochaska, J.J., Sallis, J.F., Rupp, J. and Wade, D. (2000). "Brief Self Reported Measures for Assessing Fat, Fruit and Vegetable and Calcium Intake with Adolescents." *Annals of Behavioral Medicine* 22: 132.

Proctor, M.H., Moore, L.L., Gao, D., Cupples, L.A., Bradlee, M.L., Hood, M.Y. and Ellison R.C. (2003). "Television Viewing and Change in Body Fat from Preschool to Early

Adolescence: the Framingham Children's Study." *International Journal of Obesity* 27: 827-833.

Röthlosberger, C., Calmonte, R. and Seiler, R. (1997). "Sport Stress und emotionaler Rückhalt als Determinanten von Gesundheits- und Lebenszufriedenheit bei Adoleszenten." *Psychologie und Sport* 4: 92-101

Rowlands, A.V., Ingledew, D.K. and Eston, R.G. (2000). "The Effect of Type of Physical Activity Measure on the Relationship Between Body Fatness and Habitual Physical Activity in Children: A Meta-Analysis." *Annals of Human Biology* 27: 479-97.

Sallis, J.F., Prochaska, J.J. and Taylor, W.C. (2000). "A Review of Correlates of Physical Activity of Children and Adolescents." *Medicine & Science in Sports &Exercise* 32(5): 963-975.

Schumann, A., Estabrooks, P.A., Nigg, C.R. and Hill, J. (2003). "Validation of the Stages of Change with Mild, Moderate, and Strenuous Physical Activity Behavior, Intentions, and Self-efficacy." *International Journal of Sports Medicine* 5: 363-365.

Schumann, A., Nigg, C.R., Rossi, J.S., Jordan,P.J., Norman, G.J., Garber, C.E., Riebe, D. and Benisovich, S.V. (2002). "Construct Validity of the Stages of Change of Exercise Adoption for Different Intensities of Physical Activity in Four Samples of Differing Age Groups." *American Journal of Health Promotion* 16(5): 280-287.

Sherry, B. (2005) "Food Behaviors and Other Strategies to Prevent and Treat Pediatric Overweight." *International Journal of Obesity* 29: 116-126.

Strong, W.B., Mauna, R.M., Bumkie, C.J.R., Daniels, S.R., Dishman, R.K., Guting, B., Hergenroeder, A.C., Must, A., Nicon, P.A., Pivarnik, J.M., Rosland, T., Trost, S. and Trudaeau F. (2005) "Evidence Based Physical Activity for School-age Youth." *Journal of Pediatrics* 146: 732-737.

Tucker, M. and Reicks, M. (2002). "Exercise as a Gateway Behavior for Healthful Eating among Older Adults: An Exploratory Study." *Journal of Nutrition Education & Behavior* 34(2): S14.

Ussher, M.H., Taylor, A. and Faulkner, G. (2008). "Exercise Interventions for Smoking Cessation." *Cochrane Database of Systematic Review* 4.

Vandelanotte, C., Reeves, M.M., Brug, J. and De Bourdeaudhuij, I. (2008). "A Randomized Trial of Sequential and Simultaneous Multiple Behavior Change Interventions for Physical Activity and Fat Intake." *Preventive Medicine* 46(3), 232-237.

Warren, J.M., Henry, C.J., Lightowler, H.J., Bradshaw, S.M. and Perwaiz, S. (2003). "Evaluation of a Pilot School Programme Aimed at the Prevention of Obesity in Children." *Health Promotion International* 18(4): 287-296.

Wilcox, S., King, A.C., Castro, C. and Bortz Ii, W. (2000). "Do Changes in Physical Activity Lead to Dietary Changes in Middle and Old Age?"g *American Journal of Preventive Medicine,* 18(4): 276-283.

Xu, K.T. (2002). "Compensating Behaviors, Regret, and Heterogeneity in the Dynamics of Smoking Behavior." *Social Science Medicine* 54: 133-146.

9

U.S. Food Policy and Obesity

Julian M. Alston[1], Abigail M. Okrent[2] and Joanna C. Parks[1]
[1]*Department of Agricultural and Resource Economics, University of California, Davis, CA,*
[2]*USDA Economic Research Service, Washington, DC,*
USA

1. Introduction

The obesity epidemic has been identified as the most critical public health issue facing the United States today, but it is not confined to the United States or even to high-income countries. It is a global phenomenon that reaches the entire spectrum of the income distribution, and particularly the poorest individuals within rich countries and the middle- and high-income individuals in the poorest countries. Many policies have been proposed to counter obesity, and some of those proposed policies focus on altering the food system—to influence food consumption habits and thus nutrition and obesity by changing the choices available to consumers or by changing the incentives to choose. Indeed, some jurisdictions have already introduced policies restricting the sale of certain food items in schools and others have introduced taxes on certain caloric beverages. This chapter reviews what is known about the causal links between food policies and obesity and presents new evidence about the likely efficiency and effectiveness of particular proposed policies as remedies for obesity.

We begin with a discussion of the economic rationale for government intervention in the economy to reduce the prevalence of obesity. While we note the 'public health' rationale and related arguments and instruments, our work is concerned with economic arguments and policies that work through the market for food. The economic rationale for obesity policy is based on the presence of externalities in health care (public and private) (e.g., Bhattchayra and Sood 2011; Finkelstein, Fiebelkorn, and Wang 2003; Finkelstein et al. 2009), and myopic preferences whereby individuals discount utility in distant future periods at a higher rate than in the near term (Cutler, Glaeser and Shapiro 2003; Freebairn 2010). We conclude that a basis for some such intervention exists, beyond paternalism, because some of the costs of one person's obesity are borne by others through the public health-care system. The issue then, for economists, is what is the appropriate form of market intervention? Using food policies as obesity policy is an inherently 'second-best' approach because the economic distortion does not stem from a distortion in the price of food. For example, obese individuals may create a negative externality for non-obese individuals through health insurance because of pooling of heterogeneous risk groups. In this case, a 'first-best' solution may be to charge obese individuals greater health care premiums. However, it seems likely that many people would find such a solution unacceptable, making it politically infeasible. Hence, some current policies addressing rising obesity rates have targeted changing the inputs to obesity (i.e., food consumption and physical activity).

The next major part of the chapter reviews and assesses evidence on the effects of current food policies that, according to some writers, have contributed to increasing or reducing obesity. Included in this set of policies are farm subsidies, agricultural research and development (R&D), public food and nutrition assistance programs, nutrition and health education programs, and regulations. We present evidence to show that farm subsidies have had negligible impacts on obesity (if anything, the net effect of farm support has been to increase the cost of food and thus reduce obesity), and that the Food Stamp Program (a large food assistance program), likewise, has had insignificant effects, on obesity rates. On the other hand, agricultural R&D has contributed to significant movements in the relative prices of food commodities and has most likely contributed to obesity while also yielding great benefits to society through reducing food costs for both rich and poor, thin and fat alike, and reducing pressure on natural resource stocks. In addition, we review a significant literature on the effects of nutrition labels on food for consumption in the home or away from home, the implications of changing nutrition education in schools, and the potential impacts of restrictions on what may be sold in schools. We present an up-to-date assessment of the main findings from this contemporary literature.

The third main part of this chapter presents an assessment of the likely impacts and potential usefulness of a long list of policy instruments that have been proposed—or, in some cases, adopted—to reduce the prevalence and social costs of obesity. This assessment consists of a synthesis of results from the literature and our own work in the area. One set of policies to be considered are taxes on particular foods, farm commodities as ingredients of food (e.g., sugar), or nutrients (e.g., trans fats) that affect obesity. Such policies generally are found to have small effects on food consumption, obesity and overall nutrition and health; to be regressive; and to enhance government revenue but sometimes with large social welfare costs (e.g., Kuchler, Tegene, and Harris 2004; Chouinard et al. 2007). However, Okrent and Alston (2012) found that taxes on calories would be comparatively efficient as a means of reducing obesity, and would yield significant net social benefits. Some have proposed changes to food and nutrition programs, to limit the use of food stamps (SNAP benefits) to certain types of 'healthy' foods, but economic arguments suggest that such changes may introduce more problems than they would solve (Alston et al. 2009). Other policies have been suggested, including some directed to reducing food deserts (Ver Ploeg et al. 2009) and others to do with banning advertising of fast food or otherwise regulating the food industry. These policies are also discussed briefly.

The chapter concludes with a brief synopsis of the main issues addressed and some central findings regarding the potential roles for food policy as obesity policy, which food policy instruments are potentially worthwhile, and how well they compare with other types of policies directed towards reducing obesity. Our presentation and evidence is centered on the United States, and specific to U.S. policies and institutions, but the general arguments and findings are more broadly applicable.

2. The rationale for government intervention to address obesity

The existence of large social costs of obesity alone does not justify any response by the government (Philipson and Posner 1999; Philipson and Posner 2003; Cawley 2004). Economic justifications for policies aimed at reducing obesity could rest on the existence of externalities or other economic distortions that mean the costs of being obese are not all borne by the obese individual. For instance, heavier people take up more space and impose costs on others who have to share space with them in planes, trains, automobiles, and elevators. Absent appropriate pricing policies, these costs are shared without compensation. Perhaps more serious is the phenomenon of pooling health-care system costs, both through private insurance and through Medicare and Medicaid. However, such cost pooling alone might not involve significant distortions in behavior or in total costs of obesity, and therefore it might not justify intervention by the government on economic efficiency grounds. Bhattacharrya and Sood (2011) found that the extent of moral hazard in this context—whereby the fact that costs of obesity are pooled induces responses that result in greater social costs of obesity—is quite modest. Furthermore, Bhattacharya and Sood (2011) found that the incremental health-care costs associated with obesity are passed on to obese workers with employer-sponsored health insurance in the form of lower cash wages. If this is true, some other justification for intervention is required.

Freebairn (2010) proposed two other sources of spillover effects of obesity that could be used to provide a public-goods argument to justify government intervention. First, some health-care costs are borne by government expenditure, and the use of general taxation measures to raise revenues to finance such expenditures entails deadweight losses (mainly from distortions in the labor market) such that the marginal social cost of government spending is likely to be in the range of $1.20 per dollar. Second, people who are obese are less productive than others, and have more days lost to illness, and consequently contribute less in income taxation to the total pool of government revenue available for spending on public goods. Parks, Alston, and Okrent (2011) estimated that a one pound per person increase in average adult body weight in the United States would add $749 million to annual U.S. public health expenditure. This is a lower bound estimate of the marginal social cost of the obesity externality.

A different kind of public goods argument for intervention relates to the economics of information: the government could play a role in the provision of information about the health consequences of diet to the extent that the private sector does not have sufficient incentive to do so, or in the design of appropriate regulations over the labeling of products with respect to their nutritional characteristics. The appropriate place to draw the line in such roles is far from clear, given that the private sector has some incentive to provide information that consumers demand.

To some extent at least, the arguments for government intervention related to obesity rest on paternalism—that individuals do not know what is best for them, or are unable for some reason or other to act in ways that are in their best interest, and that the state can help them make happier choices. To many people, the notion that the government could play a role in individual consumption choices—even with respect to the nutrition of infants and children—may be anathema. But this is a complex and difficult issue, an area of life in which many people are clearly unhappy with the ultimate cumulative and enduring consequences of their individual consumption choices, presumably made freely and willingly.

The psychological and the biophysical linkages between food consumption, other behavior, and obesity are complicated and dynamic, and not fully understood. To some extent, a propensity for obesity can be inherited genetically. But also, as with addictive substances like nicotine and alcohol, a propensity for obesity can be acquired through experience, beginning in infancy, or even in the womb. Such dynamic complexities, in which the daily choices made by parents can have lifelong implications for the opportunities faced by their children, are not confined to obesity; and, like other elements of child health and education, this aspect of the obesity problem may be seen by some people as a justification for policy intervention.

Obesity policy is easily justified if the policy results in a Pareto improvement in societal well-being, i.e., if the policy makes some people better off and leaves everyone at least as well off as they were before the implementation of the policy. A less stringent criterion is the Kaldor-Hicks criterion, of a potential Pareto improvement: a case where some people are made worse off by a policy change but the beneficiaries could afford to compensate the losers and still be better off. It is referred to as a potential Pareto improvement in that the compensation need not take place. This criterion for a net national benefit from a policy change is implicit in most economic discussions. Policies that effectively reduce the rate of obesity in the population would improve the health of the individuals who lose weight, benefiting them directly, and at the same time would reduce the external costs they impose on others. Such obesity policy is worthwhile, according to the Kaldor-Hicks criterion, so long as the societal benefits outweigh the costs (Deaton 2002; Just, Hueth, and Schmitz 2004, p. 32).

An alternative rationale for government intervention is the 'public health' rationale: health is a public good and the mission of public health is "fulfilling society's interest in assuring conditions in which people can be healthy" (Institute of Medicine 1988). The philosophy of social justice forms the foundation of the public health mission, that is, the public health system aims to overcome the societal barriers that prevent the equal distribution of health burdens and benefits across the population (Turnock 2004). Public health tools include the "Let's Move" campaign, increased monitoring of obesity prevalence, and stricter school food regulations. Given that low-income individuals in the United States experience more extreme obesity (BMI > 35) than higher-income individuals, and that some of this disparity may be attributable to the societal conditions associated with living in poverty (e.g., lack of health insurance and access to fresh fruits and vegetables), there is a social justice or public health argument for public policy to reduce obesity (Jolliffe 2010).[1] While we acknowledge this alternative rationale for policy, in this chapter we focus on policies that operate through markets for food, for which the relevant rationale must be an economic one.

3. Effects of past and current food policies on obesity

The increased prevalence of obesity in the United States has been attributed to past and current policies directed at both producers and consumers of food. Policies directed at improving the welfare of farmers (i.e., farm subsidies and investment in agricultural R&D) and low-income families (i.e., food and nutrition assistance programs) may have inadvertently contributed to increased consumption of food, calories and body weight. In addition, policies directed at providing better health and nutrition information (including food labels) to consumers so as to help them make better-informed consumption decisions, may have been ineffective. This section examines the literature on the effects of past and current food policies on obesity, with an emphasis on U.S. policies.

---

[1] In the United States the most common definition of low-income is an individual or household that has an income at or below 130 percent of the federal poverty line.

3.1 Farm subsidies

The United States has a long history of agricultural policy and many commentators—including prominent economists, nutritionists, journalists, and politicians—have claimed that American farm subsidies have contributed significantly to the 'obesity epidemic.' They argue that farm subsidies have made fattening foods relatively cheap and abundant, and that reducing these subsidies will go a long way towards solving the problem. These commentators often treat the point as self-evident, and do not present details on the mechanism by which farm subsidies are supposed to affect obesity, nor evidence about the size of the likely impact. In particular, Pollan (2003, 2007) has claimed that subsidies on commodities such as corn and wheat have led to lower prices of high-calorie, processed foods. As proof of this effect, Pollan has pointed to the correlation between increased subsidies to corn farmers and rising obesity rates in the United States between 1970 and 2005. Likewise, Nestle (2002), Tillotson (2004), Muller, Schoonover and Wallinga (2007), Ludwig and Pollak (2009) and Popkin (2010) have attributed the growth in U.S. obesity rates to agricultural policies, and advocated a reorientation of government spending away from corn and wheat to fruits, vegetables and whole grains.

Farm commodity programs can affect the rate of obesity by changing the relative prices of food commodities and thus retail foods, and hence, the composition of food consumption. However, several studies have demonstrated that the magnitude of this effect in the United States is likely to have been small and ambiguous (Alston, Sumner, and Vosti 2006; Alston, Sumner, and Vosti 2008; Beghin and Jensen 2008; Miller and Coble 2007; Rickard, Okrent, and Alston 2011; Schmidhuber 2004; Senauer and Gemma 2006). This finding has several elements.

It is true that farm subsidies have resulted in lower U.S. prices of some commodities such as food and feed grains, and consequently, lower costs of producing cereals and bakery products and meats. However, the price-depressing effects of subsidies has been contained (or even reversed) by the imposition of additional policies like acreage set-asides that restrict production. In addition, since 1996, about half of the total subsidy payments have become decoupled from production and based on historical rather than current acreage and yields (Alston, Sumner and Vosti 2008; Sumner 2005; Beghin and Jensen 2008). Reflecting these facts about the policies, complete elimination of U.S. commodity subsidies would have minimal effect on corn, wheat or rice production and hence prices: 9–10 percent decrease in price of corn, 4–6 percent decrease in price of rice and 6–8 percent decrease in price of rice (Sumner 2005).

Conversely, some farm commodity programs have actually increased the prices of commodities. Trade barriers on sugar, dairy and orange juice have increased the cost of these commodities to U.S. buyers and the U.S. food industry. The combination of subsidies for some commodities and trade barriers for others makes the story complicated at times. A case in point is the market for caloric sweeteners. Corn is often the target of criticism as a contributor to obesity, especially because of its use to make high fructose corn syrup (HFCS), which is used as a caloric sweetener in many foods and beverages. The use of HFCS as a sweetener has been encouraged by U.S. sugar policy that made sugar much more expensive and gave food manufacturers an economic incentive to substitute HFCS for sugar. So, farm subsidies are responsible for the growth in the use of corn to produce HFCS as a caloric sweetener, but not in the way that is often suggested. The culprit here is not corn subsidies; rather, it is sugar policy that has restricted imports, driven up the U.S. price of sugar, and encouraged the replacement of sugar with alternative caloric sweeteners. Combining the sugar policy with the corn policy, the net effect of farm subsidies has been to increase the price of caloric sweeteners generally, and to discourage total consumption while causing a shift within the category between sugar and HFCS. In this context, the subsidy policies effectively increase the overall price of caloric sweeteners; removing policies applied to sweeteners would lead to an overall increase in consumption of sweeteners (Alston, Rickard, and Okrent 2010).

3.2 Agricultural R&D

Other agricultural policies may have had more significant effects on obesity. Alston, Sumner and Vosti (2006, 2008) suggested that productivity gains resulting from agricultural research and development (R&D) have been much more important than commodity subsidies as a determinant of food prices. In real terms agricultural commodity prices trended down significantly during the past 50 years, reflecting growth in supply of agricultural products outstripping growth in demand that was fueled by increases in population and per capita incomes. Alston, Beddow and Pardey (2009) attributed these trends in prices primarily to growth in farm productivity, which they ascribed primarily to public and private investments in agricultural R&D. Likewise, Miller and Coble (2007, 2008) estimated that increases in total factor productivity contributed more to lowering prices of retail food products, and thus, the portion of income spent on food, than did subsidies to farmers in the United States and across OECD countries. Beghin and Jensen (2008) also attributed substantial declines in the price of corn, and hence, HFCS and food products that use HFCS to technical change rather than subsidies.

3.3 Public food and nutrition programs

Several studies have investigated the effects of participation in Food and Nutrition Programs (FANPs) on obesity (e.g., see Ver Ploeg, Mancino, and Lin 2007). This section emphasizes the three main programs: the *Food Stamp Program* (FSP), which in 2008 was revised and renamed the *Supplemental Nutrition Assistance Program* (SNAP); the *Special Supplemental Nutrition Program for Women, Infants, and Children* (WIC); and the *National School Lunch Program* (NSLP).[2]

The FSP has been much studied, with much of the research based on analysis of data from large national surveys, and mainly concerning impacts on household food expenditures, household nutrient availability, and individual dietary intakes. The research suggests that, participation in the FSP increases household availability of energy and protein and perhaps some vitamins and minerals. Less evidence is available about the impacts on individuals' dietary intakes. Fox, Hamilton, and Lin (2004a, b) concluded that results were mixed and

---

[2] While the FSP (or SNAP), WIC, and NSLP are by far the largest and most widely available FANPs, other FANPs are important for different subsets of the population. For example, the School Breakfast Program (SBP) serves about 10.6 million children per day, compared to the NSLP's 31.3 million. The Child and Adult Care Food Program (CACFP) serves about 1,831.1 million meals a year to children in day care, and 64.2 million meals a year to adults, with about 3.3 million participants a day receiving a meal. Together, the SBP and the CACFP cost $5.1 billion per year, compared with $10 billion for the NSLP in 2009. The other seven smaller programs together cost only $1.2 billion annually.

collectively inconclusive concerning impacts of the FSP on several other nutrition- and health-related outcomes (such as birth weight, height, weight, nutritional biochemistries, and general measures of health status). Although the findings are mixed, the balance of the evidence (weighing those studies that have done a better job of addressing the perennial econometric challenges) indicates that women who participate in the FSP are more likely to be obese, with estimated probabilities ranging from 6 to 9 percent (Gibson 2003; Chen, Yen and Eastwood 2005; Meyerhoefer and Pylypchuk 2008). However recent work by Fan (2010) found no effect of long- or short-run FSP participation by women on obesity once pre-participation weight and other socio-economic and demographic factors were taken into account. More-recent research by Parks (2011) demonstrated that low-income women who participated in the FSP at some point in the previous year did not gain significantly more weight over the past year then eligible women who did not participate.

Studies of the impact of WIC have emphasized impacts on health-and nutrition-related outcomes of participants, which is the primary goal of WIC, unlike the other FANPs. Among others, obesity and abnormal gestational weight gain are both considered nutritional risk factors that qualify women for participation in WIC. The available evidence suggests that WIC participation increases the intakes by pregnant women of most of the target nutrients; less clear is whether it has led to a greater prevalence of adequate intakes, and little evidence is available on whether the dietary intakes of WIC participants are more-closely aligned with the *Dietary Guidelines for Americans*. Likewise, very little compelling evidence is available on the impacts of WIC on dietary intakes of participating children or on their nutrition and health characteristics. However, a range of evidence indicates that WIC has had significant impacts on dietary intake of infants, including particular nutrients that are largely associated with the consumption of cow's milk versus formula and the introduction of solid foods. In particular, WIC has significantly reduced the prevalence of anemia among low-income American children. Children from households with WIC participants also tended to have better general health status, more so for the lowest-income children, and a higher probability of having up-to-date immunizations.

Some writers have suggested that the NSLP and other school meals programs may have contributed to rising rates for obesity among school children.[3] Ralston, Newman, Clauson, Guthrie and Buzby (2008) provided a comprehensive discussion of the NSLP, including historical trends, participant characteristics, and challenges facing administrators of the program, including tradeoffs between nutritional quality of foods served, costs, and participation—which they refer to as a school meals 'trilemma'—as well as between program access and program integrity. The authors reported that program participation has had little if any measurable impact on caloric intake or obesity; that participants derive important nutritional benefits from participating in the program, including higher intake of key nutrients and under-consumed foods and lower intake of sweets, but also have high intakes of fat and sodium. A significant part of the problem of nutritional quality is associated with costs. Many schools depend on revenues from 'competitive foods,' even though such foods have been found to contribute to overconsumption of calories, increased plate waste of nutritionally balanced NSLP lunches, and decreased intakes of nutrients by students.

---

[3] Other factors contributing to obesity related to eating at school include the availability of dispensers for candy, sodas, or fast foods, and the proximity of fast-food restaurants.

3.4 Nutrition and health education (including food labels)

Consumers make decisions on what foods to purchase based on information on prices and attributes of foods (e.g., convenience, healthiness, and so on). If individuals do not have a clear understanding of the health attributes of foods and how foods affect body weight, or they are unaware of the connection between obesity and higher risks of chronic illnesses, then individuals may be inadvertently choosing foods that cause their weight to deviate from the medically ideal body weight. Two types of information on diet and nutrition are relevant: (i) information on the negative health aspects of obesity, and (ii) information on how food consumption translates into personal weight gain or loss. In the United States, studies find that many individuals do not have a clear understanding of the diet-disease connection. Some evidence shows that increases in individual diet-disease knowledge significantly decrease the probability that an individual is obese and suggests that new policies to increase diet-disease health knowledge should lead to decreases in the incidence of obesity (Nayga 2000; Falba and Busch 2005). For example, Nayga (2000) estimated that complete acquisition of diet-disease knowledge could reduce the likelihood of an individual being obese to roughly 20%. This estimate of the impact of information is large and most likely overstated because of endogeneity between diet-disease knowledge and BMI: individuals who are more knowledgeable about the connection between diet and risk of disease are more likely to be those who have made weight management a priority for reasons that are unobserved, and not controlled for in estimation.

The second type of information concerns individual understanding of the translation from current food consumption into future weight outcomes. Nutritional labeling increases the ability of an individual to predict the effect of food consumption on future weight. In the United States, many studies that examine the effect of nutritional labels on grocery store purchases utilize variation between 1992 and 1999, before and after the passage of the National Labeling and Education Act (NLEA). Labels existed before the NLEA under the voluntary labeling rules established by the Food and Drug Administration (FDA) in 1975, but they were not on all packaged foods and were not standardized. Mojduszka and Caswell (2000) examined labeling information on packages in 33 food product categories and concluded that incentives for voluntary disclosure of nutritional content by food processing firms prior to the NLEA did not generally result in reliable and consistent nutrition information being made available to consumers. Mathios (2000) found that prior to the NLEA, all low-fat salad dressings had a nutrition label, while the majority of higher-fat dressings did not, and sales for those with the highest fat levels declined significantly after the NLEA. Variyam and Cawley (2007) compared the change in body weight, after the implementation of NLEA, among those who use labels when food shopping to that among those who do not use labels. They found that non-Hispanic white women benefited the most from nutritional labeling, with an estimated 3.36 percent reduction in obesity associated with label use, whereas the new labels had no effect on the body weight of white men and black women and were actually associated with an increase in the body weight of black men.

Variyam (2005) noted that the NLEA exempts much of the food consumed at restaurants from mandatory labeling regulations. Because consumers are less likely to be aware of the ingredients and nutrient content of restaurant foods than of foods prepared at home, public health advocates have called for mandatory nutrition labeling for major sources of these foods, such as fast-food restaurants and chain restaurants. Recent studies have begun to examine the effect of mandatory calorie postings on restaurant menus in New York City. Downs et al.

(2009) utilized receipts collected from patrons outside two burger restaurants and a coffee shop in New York City before and after the mandate, and found little evidence of any effect of calorie postings on consumption. Elbel et al. (2009) also utilized receipts collected from patrons outside 14 chain restaurants in New York City and five chain restaurants in Newark, New Jersey. Similar to Downs, Loewenstein and Wisdom (2009) they found little evidence to suggest that the labels had any effect on consumption. However, Bollinger, Leslie and Sorensen (2010) found that mandatory calorie posting influenced consumer behavior at Starbucks, resulting in a 6% decrease in the average calories per transaction (down from 247 to 232 calories per transaction). Almost all of the observed effect was related to food purchases, where average food calories per transaction fell by 14%. Interestingly, they also found that calorie postings did not result in any statistically significant change in Starbucks' total revenue.

3.5 Regulation

The majority of evidence indicates that unregulated food marketing to children has contributed to the obesity problem. Concern has been expressed about the consequences of marketing food and other goods to children since the 1970s, but no progress has been made in establishing regulations that the government would enforce. Rather, the United States has relied on the food industry to regulate itself (Wilde 2009). Children under the age of approximately 8 (and possibly as old as 11) are especially vulnerable to food advertising because they cannot distinguish the content in a television program from the content in a commercial, or comprehend that the purpose of an advertisement is to persuade. The research summarized by the American Academy of Pediatrics (2006) and the Institute of Medicine (IOM)(2005) showed that television advertising—which only accounts for a fraction of the total advertising children are exposed to—strongly affects food preferences, short-term consumption (versus 'usual dietary intake'), and purchase requests in 2–11 year olds. Some evidence suggests that advertising affects beliefs about foods and beverages and usual dietary intakes among children 2–11 and 2–5 years of age, respectively (American Academy of Pediatrics 2006; IOM 2005)[4]. In 2005, the IOM called on the U.S. food industry to self-regulate and reduce the prominence of energy-dense and nutrient-poor foods in advertisements aimed at children. However, the IOM also noted that, if self-regulation failed to achieve this goal, Congress should step in and enact policy that would mandate the changes in marketing to children (Institute of Medicine 2005; Wilde 2009).

4. Proposed and potential food policies

Rising (or just high) rates of obesity, especially among children, have attracted the attention of governments worldwide. The issue is significant in the United States. Food and nutrition policy that was once devoted to issues of food security and inadequate nutrition must now confront the modern malnutrition paradox: poor people having poor dietary quality and food insecurity while at the same time experiencing the health consequences of overeating and sedentary lifestyles. Many proposals for policies to address this situation have been

---

[4] Usual dietary intake refers to the long-term dietary intake patterns of an individual, i.e., the average daily intakes of different dietary components over a year. Short-term consumption refers to consumption during a specific time frame following a particular stimulus or event, e.g., the number of potato chips and apple slices consumed after exposure to a potato chip advertisement (IOM 2005).

suggested, including changes to existing food and nutrition programs, as well as other policies that also work through changing the effective prices or availability of food generally, or particular foods. Some such policies have already been introduced in some places (e.g., 35 states had a sales tax on soda in 2011).[5]

4.1 Food taxes and subsidies

Economists have modeled and measured the potential impacts of various taxes and subsidies as instruments for reducing obesity. Some studies suggest that taxation or subsidization of certain foods would be effective as means of reducing average body weight in the United States and in other developed countries (Smith, Lin, and Lee 2010; O'Donoghue and Rabin 2006; Cash, Sunding, and Zilberman 2005; Sacks et al. 2011). A tax on a class of foods that are energy dense and deemed 'unhealthy' (e.g., soda and chips) would make 'unhealthy' foods more expensive relative to 'healthy' foods such that consumers would substitute away from consumption of 'unhealthy' foods and into consumption of 'healthy' foods. Others have argued that such pricing policies would have little effect on food consumption, and hence obesity and may also be regressive (Schroeter, Lusk, and Tyner 2007; Kuchler, Tegene and Harris 2004; Gelbach, Klick, and Strattman 2007; Allais, Bertail, and Nichèle 2010). It has also been suggested that even if a tax on a particular food is ineffective at reducing consumption of 'unhealthy' foods, the tax revenues that are generated from the tax could be used to fund public information programs and other obesity-reducing strategies (Jacobson and Brownell 2000; Kuchler, Tegene, and Harris 2005; Brownell and Frieden 2009).

An alternative to taxing a particular 'unhealthy' food is subsidizing a food deemed to be 'healthy.' Many nutritionists recommend eating more fruits and vegetables as a weight-control strategy because fruits and vegetables are low-calorie, high-fiber foods that have been found to be more filling and satisfying than low-fiber foods (Tohill 2004). Guthrie (2004) and Lin and Guthrie (2007) argued that policies that make retail fruits and vegetables products cheaper would help reduce obesity by causing consumers to substitute away from more energy-dense foods. Evidence in support of this argument is mixed. While fruit and vegetable subsidies may cause individuals to consume more healthfully, they may also consume more calories not only from fruits and vegetables, but from goods that are complements to fruits and vegetables (Okrent and Alston 2012).

Rather than taxing a food group or product, a tax on a nutrient (i.e., total fat or saturated fat), on an ingredient (e.g. added-sugars), or on the energy content of a food could be used to address obesity. Chouinard et al. (2007) estimated that taxing the fat content of dairy products by 10 percent in the United States would decrease fat consumption but would lead to a reduction in weight by less than one pound per person per year, holding all other determinants of weight constant. They also found that the consumer welfare losses from the fat tax on dairy products would be slightly more than the revenue generated (–$4.48 billion versus $4.45 billion) and the tax would be highly regressive. Miao, Beghin and Jensen (2010) reported similar findings for taxing the sugar content of foods.

Okrent and Alston (2012) developed a model of the U.S. farm and food industry expressly designed for analyzing such questions. Key findings from this work are in keeping with

---

[5] Information on soda taxes is taken from the Bridging the Gap Program at the University of Illinois at Chicago; available at: http://www.impacteen.org/obesitystatedata.htm#01.

economic intuition. Given that obesity is caused by an energy imbalance—an excess of calories consumed over calories expended—a tax on food according to its caloric content is likely to be a relatively efficient instrument (having the lowest total social cost per unit of impact on obesity). Taxes applied to particular nutrients (e.g., taxes on food products based on their sugar content or fat content) are likely to be less efficient than a tax on calories; taxes on particular foods (e.g., caloric beverages or sodas) are likely to be even less efficient. Many such instruments are likely also to be regressive, in the sense that the burden will be borne disproportionately by the poor.

Economic tools like 'fat taxes' and 'thin subsidies' could be used both to influence consumption habits and raise revenue to offset excessive medical public medical expenditures engendered by obesity. Small taxes and subsidies have been found to have little effect on the consumption of categories of foods deemed 'unhealthy' and the overall caloric content of foods. Even if such taxes or subsidies were enacted, at what level of the food processing chain should they be applied? Would a subsidy for research and development into farm production of fruits and vegetables lower the price of those commodities and thereby encourage substitution from unhealthy foods to 'healthy' foods? Or would a tax on fast food be more effective?

4.2 Changes to farm subsidies or the emphasis of agricultural R&D

Many people blame federal farm subsidies for the current obesity problems (e.g., Pollan 2003, 2007, 2008). It may seem obvious that subsidies must make certain foods cheaper, therefore contributing to overconsumption, but U.S. farm policies have had generally modest and mixed effects on prices and quantities of farm commodities. The overall effect on the prices paid by U.S. consumers for food has been negligible and, consequently, eliminating farm policies would have a tiny influence on dietary patterns and obesity (Alston, Sumner and Vosti 2008; Miller and Coble 2008; Alston, Rickard and Okrent 2010; Rickard, Okrent, and Alston 2012). U.S. farm policies might well be seen as unfair and inefficient. But whether we like these policies or not for other reasons, their effects on obesity are negligible. In fact, eliminating all farm subsidies, including those provided indirectly by trade barriers, may, if anything, lead to an *increase* in annual per capita consumption of calories and an increase in body weight.[6]

In contrast, agricultural R&D has had substantial impact on the abundance and relative prices of farm food commodities. Since 1949, the overall price of farm commodities in the United States has fallen by over 60 percent relative to the GDP deflator. This decline in the real prices of food commodities is attributable largely to agricultural productivity growth, mostly due to agricultural R&D, a significant share of which was funded (and in some cases performed) by the U.S. government, mainly through the USDA. If cheap and abundant food has contributed to obesity, then R&D from the USDA is partly to blame. Does this mean the USDA should have done less research or different research? The evidence from studies of

---

[6] Similar results have been found for the case of sugar policy in the European Union. The EU uses a combination of price floors, import duties, export subsidies and quotas to support domestic sugar farmers. The elimination of these policies would reduce the price of sugar by 36% over a 4-year period. Bonnet and Requillart (2010) found that this 36% reduction in price in France would increase consumption of regular soft drinks by 1 liter per year per person and the consumption of added sugar by 124 grams per person per year. See, also, Schmidhuber (2004).

the returns to research indicates that the United States has persistently underinvested in agricultural R&D. Marginal benefit-cost ratios in the range of 20:1 and higher indicate that the extent of the underinvestment has been significant (e.g., see Alston, Andersen, James, and Pardey 2010). These measures did not account for the contributions of agricultural R&D to the social cost of obesity, but it does not seem likely that taking those costs into account would change the picture appreciably. Nor does it seem likely that taking the impacts on costs of obesity into account would imply a significant change in the emphasis of the national benefit-maximizing research portfolio, to increase the share of funding going to research on specialty crops (e.g., see Alston and Pardey 2008). It is likely that better-targeted instruments can be found for reducing social costs of obesity, rather than reducing support for agricultural R&D generally (which is already underfunded) or sharply changing the mix to reduce research on commodities such as livestock and grains in favor of 'healthy' foods like fruits and vegetables (a change that might substantially reduce the returns to the portfolio as a whole without much appreciable impact on obesity). But the precise terms of this tradeoff are not clear and further work is needed before we can make more-specific claims.

4.3 Changes to food and nutrition assistance programs

With respect to both the FSP and the NSLP discussions have emphasized the role of the cost advantages of foods and ingredients that are relatively energy dense in highly budget-constrained dietary choices. Revisions to the programs have been suggested with a view to disallowing certain types of foods. Meanwhile, the question remains somewhat open as to what is the size of the effect, if any, of program participation on obesity rates, what is the social cost of that effect, which should be charged against the programs, and what adjustments to the programs may be appropriate to reduce those costs and thereby enhance the efficiency of the programs. In light of the observation that many of the poor are also obese, it has been suggested that the FSP could be modified to encourage participants to eat healthier diets. This idea may have been encouraged by findings that the FSP may have contributed to obesity among participants, but does not rest on that possibility.

Under the FSP, participants may redeem their coupons for almost any food items. Some have suggested changing the program to restrict food stamp purchases to exclude certain foods deemed to be 'unhealthy,' e.g., sugar sweetened beverages (Brownell and Frieden 2009; Brownell and Ludwig 2011). Economic analysis (e.g., Guthrie et al. 2007; Alston, Mullally, Sumner, Townsend and Vosti 2009) suggests such restrictions would be ineffective or counterproductive because they would (i) have no impact on food consumption choices by many participants, (ii) would discourage program participation by some others, (iii) if effective in changing consumption choices, would result in relative price changes that would have opposite effects on some groups, and (iv) would increase administrative costs of the program as the food industry would redesign foods to meet the criteria in the restrictions. Another proposal would give coupons greater purchasing power if applied to particular categories of foods (e.g., fresh fruits and vegetables) than others (see Guthrie et al. 2007). Such a modification may be more effective than restricting the list of items eligible for purchase with food stamps, though it would have some of the same drawbacks. Both of these ideas would use a single instrument, the FSP, to pursue two targets: (i) assuring adequate nutrition intakes in populations deemed at risk of under-nutrition (to be achieved by an expanded food budget constraint for the poor); and (ii) reducing the prevalence of obesity. Modifying the FSP to pursue the second objective would almost surely diminish its effectiveness in relation to the first.[7] Moreover, the problem of obesity is not confined to food stamp participants, and other instruments will be necessary if the government means to address the broader problem. If such other instruments are applied more broadly, it may make less sense to modify the FSP as well.

Congress regularly reauthorizes federal school meal and child nutrition programs. However, the Healthy, Hunger-Free Kids Act of 2010 includes several notable changes in child meal program nutritional standards, access, and funding, as well increased school accountability and monitoring for meal quality. The new legislation provides the largest increase in funding and reimbursement rates for the school lunch program in decades, makes it easier for qualified children to receive free school meals, extends after-school meals to more at-risk children, and provides additional technical assistance to local school food-service providers. The Healthy, Hunger-Free Kids Act will increase access by improving and simplifying application procedures and expanding universal eligibility in high-poverty neighborhoods.

The legislation also allows for the elimination of soda and other junk foods from schools nationwide. The USDA Food and Nutrition Service proposed new Nutrition Standards for the NSLP and SBP in mid-January 2011.[8] The recommended changes to NSLP and SBP nutrition standards include increased servings of fruits, vegetables, and whole grains, a ban on trans fat, and reduced sodium content, as well as, for the first time ever, a maximum allowable calorie content per meal (the nutrition standards already include a calorie minimum). The Healthy, Hunger-Free Kids Act also raises nutritional standards for child-care centers participating in the Child and Adult Care Food Program, and provides funding for training, technical assistance and tools to assist child-care providers in complying with new standards and promoting better nutrition wellness among young children.

4.4 Changes to nutrition and health education

Several studies have found that differences in health knowledge account for some variation in current obesity rates and that the introduction of nutrition labels has been somewhat effective in changing the food consumption patterns of individuals. However, strategies aiming to improve the state of nutrition knowledge of individuals (e.g., nutrition education programs and food labelling regulations) have been relatively ineffective, and policymakers have suggested several changes.

Since the 1990 enactment of NLEA, the FDA has required packaged foods to list the amounts of various nutrients per serving along with a definition of serving size. This requirement has had some effects on consumer purchasing behavior but the actual use of the labels is much less than what is typically reported, fewer consumers are using the labels than in the past, and labels often leave shoppers confused and misinformed (Cowburn and

---

[7] Conversely, some would favor expanding the scope of the FSP to include among the eligible purchases food away from home, including fast food, at least for certain groups in society (such as homeless) as a way of better achieving its primary purpose of providing a food income safety net. Such a proposal was debated in the New York Times http://www.nytimes.com/roomfordebate/2011/09/27/expand-the-use-of-food-stamps?nl=todaysheadlines&emc=thab1

[8] The complete list of proposed nutrition standards is available at:
http://www.fns.usda.gov/cnd/governance/regulations/2011-01-13.pdf.

Stockley 2005; Todd and Variyam 2008). New forms of food labeling, and front-of-the package nutrient postings have been cited as potential tools for improving the nutrition of the population (Nestle and Jacobson 2000). In particular, many developed countries have begun to investigate whether 'traffic-light' nutrition labels with symbols on packages that indicate high, medium or low levels of a few specific nutrients or energy will help consumers make more healthful food choices. The evidence of whether traffic-light labels work is mixed. First, most of the front-of-the-package labeling has been self-regulated by the food industry—the 'Smart Choices' program in the United States, the 'Choices' logo in the Netherlands, and the 'Multiple Traffic Light' in the United Kingdom. Roberto et al. (2011) found that 60 percent of the foods that had the 'Smart Choices' label did not meet standard nutritional criteria for a healthy food based on the Nutrient Profile Model (a non-industry-developed and validated national standard).

Second, even though most of these programs have been found to help consumers identify which foods are healthy, very little evidence is available on whether the programs actually caused consumers to purchase healthier foods (Grunert and Wills 2007). Garson and Engelhard (2007) found that sales of foods in vending machines at the University of Virginia labeled with a red light (i.e., highest in fat and calories) decreased 5.3 percent, while those labeled with a yellow light or green light increased by 30.7 and 16.5 percent, respectively. Similarly, Vyth et al. (2010) found some support for the claim that the Choices logo motivated healthy food choices in the Netherlands. Alternatively, Sacks, Rayner and Swinburn (2009) found little evidence that the introduction of the Multiple Traffic Light labels in the United Kingdom on ready meals and sandwiches had any impact on the relative healthfulness of consumer purchases. Hence, even if the new labels clarify whether a food is healthy, it is not clear the consumers will be induced by the information to change behavior and eat less calories.

The Healthy, Hunger-Free Kids Act establishes important reforms for the national food stamp nutrition education program (SNAP-Ed), expanding the program's focus to include obesity prevention and allowing the use of community and public health approaches to improve the diets of low-income families with the Nutrition Education and Obesity Prevention Grant Program. The provision requires that nutrition education activities be evidence-based and focused on specific nutrition or health outcomes. It also requires that schools that participate in the NSLP or SBP implement a 'school wellness plan,' with physical activity and nutrition education goals, plus nutrition guidelines for foods available on campus, by the 2011-12 school year.

4.5 Food marketing regulations

When parental supervision declines at home, children are left vulnerable to environmental stimuli that can affect consumption patterns. For example, food marketers target the youth demographic by advertising during periods of the day when children watch television, which could induce children to choose less-healthy, but heavily advertised foods. Rashad, Grossman and Chou (2005) explained that a brand's advertising acts as a complementary good such that consumers derive more utility from consuming a more-advertised good. For example, a child may value the toys associated with commercials more than a toy that is not advertised. Their results show a strong positive effect of exposure to fast-food restaurant advertising on the BMI for children and adolescents; a complete ban on fast-food television advertising during children's programming hours would reduce the number of overweight children and adolescents by 10 percent and 12 percent, respectively.

The Omnibus Reconciliation Act of 2009 called on the Federal Trade Commission, the USDA, the Centers for Disease Control and Prevention (CDC), and the FDA to jointly create the Interagency Working Group on Food Marketed to Children (Working Group) and produce a set of standards for the advertising of foods to children and adolescents. In April 2011 the Working Group released for public comment a set of proposed industry guidelines for self-regulation. The Working Group proposed two main nutrition principles, labeled principles A and B, that foods marketed to children should meet, as well as a definition of "marketing targeted at children and adolescents." Nutrition principle A indicates that food products marketed predominantly to children (ages 2–17) should "make a meaningful contribution to a healthful diet." Nutrition principle B calls on the industry to develop and reformulate food products marketed predominantly to children and adolescents to minimize the proportion of nutrients that could have unwanted effects on body weight (e.g., added sugars, salt, and trans fats) or health of children. Lastly, the Working Group adopted the FTC definition of "marketing targeted at children and adolescents" (ages 2–17), and requested that the industry adhere to the Nutrition Principles in 20 different categories of targeted advertising and promotional activities including movies, video games, magazines, and specific websites based on the share of the audience represented by children (Interagency Working Group 2011).

5. Conclusion

Government intervention in the food industry with the aim of reducing the incidence of obesity can be justified, economically, if obesity entails externalities and if the benefits from the intervention exceed the costs. Food policy is a second-best obesity policy, but some of the policies considered in this chapter could pass this benefit-cost test, given that first-best policies are not feasible. Various policies have been proposed that would counter obesity by making relatively unhealthy foods relatively expensive. Some of the policies that are commonly discussed would be ineffective or inefficient. These include (i) limiting food stamps to only healthy foods, (ii) taxing foods according to their fat or sugar content, (iii) taxing caloric beverages, (iv) eliminating farm subsidies, or (v) reducing funding for agricultural R&D. Most of these policies would be ineffective relative to the objective of reducing obesity, and may be counterproductive relative to the safety-net purpose in that almost any such policy is likely to be regressive, falling disproportionately heavily on the poor. All of these policies would be inefficient in the sense that other policies could be devised that would have the same benefit (in terms of impact on obesity) at lower social cost.

A relatively efficient policy would be to tax foods according to their individual caloric content, but this policy too might not be very effective unless accompanied by other instruments directed towards other elements of this very complex social problem. A combination of incentives through tax and subsidy policies, selective regulation, and educational programs seems likely to be more efficient than taxes alone. And the design of such policies bears careful consideration and analysis. Discussions of obesity tax policies tend to emphasize the demand side—taking the food product choices as given, and focusing on the use of incentives for consumers to choose differently from among the given choices. But given the large and increasing role of prepared foods and food away from home, the supply side may offer greater scope for substitution away from energy-dense foods, through innovations in food processing and manufacturing. Hence, in designing any such policy it is desirable to have in mind the potential impact on the food industry, through incentives to innovate in manufacturing and produce and promote consumption of foods that are attractive to consumers while having socially desirable nutritional characteristics. Calorie taxes could be introduced based on economic efficiency alone, but they might be based partly on paternalism, which is a potentially dangerous policy path. On the other hand, FANPs themselves are paternalistic, and it seems perfectly reasonable to require that FANPs do not encourage the poor or their children to eat unhealthy foods. This argument implies a basis for adapting the programs that provide specific foods (such as WIC, NSLP, SBP) and eliminating junk foods from schools, as entailed under the Health, Hunger-Free Kids Act of 2010. It is less useful relative to the SNAP program given its generic purpose as a food safety net.

6. References

Allais, O., Bertail, P. and Nichele, V. (2010). "The Effects of a Fat Tax on French Households' Purchases: A Nutritional Approach." *American Journal of Agricultural Economics* 92(1): 228-245.

Alston, J.M. (2009). "Efficiency of Income Transfers to Farmers through Public Agricultural Research—Theory and Evidence from the United States." *American Journal of Agricultural Economics* 91(5): 1281-1288.

Alston, J.M., Andersen, M.A., James, J.S. and Pardey, P.G. (2010). *Persistence Pays: U.S. Agricultural Productivity Growth and the Benefits from Public R&D Spending*. New York: Springer.

Alston, J.M., Beddow, J.M. and Pardey, P.G. (2009). "Mendel versus Malthus: Research, Productivity and Food Prices in the Long Run." Department of Applied Economics Staff Paper No. P09-1. St. Paul, MN, University of Minnesota.

Alston, J.M., Mullally, C., Sumner, D.A., Townsend, M. and Vosti, S.A. (2009). "Likely Effects on Obesity from Proposed Changes to the U.S. Food Stamp Program." *Food Policy* 34: 176-184.

Alston, J.M. and Pardey, P.G. (2008). "Public Funding for Research into Specialty Crops." *HortScience* 43(5): 1461-1470.

Alston, J.M., Rickard, B.J. and Okrent, A.M. (2010). "Farm Policy and Obesity in the United States." *Choices* 3rd Quarter 2010.

Alston, J.M., Sumner, D.A. and Vosti, S.A. (2006). "Are Agricultural Policies Making Us Fat? Likely Links Between Agricultural Policies and Human Nutrition and Obesity, and their Policy Implications." *Review of Agricultural Economics* 28(3): 313-322.

Alston, J.M., Sumner, D.A. and Vosti, S.A. (2008). "Farm Subsidies and Obesity in the United States: National Evidence and International Comparisons." *Food Policy* 33(6): 470-479.

American Academy of Pediatrics Committee on Communications (2006). "Children, Adolescents, and Advertising." *Pediatrics* 118(6): 2563-2569.

Beghin, J. and Jensen, H. (2008). "Farm Policies and Added Sugars in US Diets." *Food Policy* 33: 480-488.

Bhattacharya, J. and Sood, N. (2011). "Who Pays for Obesity?" *Journal of Economic Perspectives* 25(1): 139-158.

Bollinger, B., Leslie, P. and Sorensen, A. (2010). "Calorie Posting in Chain Restaurants." NBER Working Paper Series No. 15648. National Bureau of Economic Research.

Bonnet, C. and Requillart, V. (2011). "Does the EU Sugar Policy Reform Increase Added Sugar Consumption? Empirical Evidence on the Soft Drink Market." *Health Economics* 20(9): 1012-1024.

Brownell, K.D. and Frieden, T. (2009). "Ounces of Prevention--The Public Policy Case for Taxes on Sugared Beverages." *The New England Journal of Medicine* 360(18): 1-4.

Brownell, K.D. and Ludwig, D.S. (2011). "The Supplemental Nutrition Assistance Program, Soda, and USDA Policy: Who Benefits?" *Journal of the American Medical Association* 306(12): 1370-1371.

Cash, S.B., Sunding, D.L. and Zilberman, D. (2005). "Fat Taxes and Thin Subsidies: Prices, Diet, and Health Outcomes." *Acta Agriculturae Scandinavica Section C 2*: 167-174.

Cawley, J. (2004). "An Economic Framework for Understanding Physical Activity and Eating Behaviors." *American Journal of Preventive Medicine* 27(3S): 117-125.

Chen, Z., Yen, S.T. and Eastwood, D.B. (2005). "Effects of Food Stamp Participation on Body Weight and Obesity." *American Journal of Agricultural Economics* 87(5): 1167-1173.

Chouinard, H.H., Davis, D.E., LaFrance, J.T. and Perloff, J.M. (2007). "Fat Taxes: Big Money for Small Change." *Forum for Health Economics & Policy* 10(2).

Cowburn, G. and Stockley, L. (2005). "Consumer Understanding and Use of Nutrition Labeling: A Systematic Review." *Public Health Nutrition* 8(1): 21-28.

Cutler, D.M., Glaser, E.L., and Shapiro, J.M. (2003). "Why Have Americans Become More Obese?" *Journal of Economic Perspectives* 17(3): 93-117.

Deaton, A. (2002). "Policy Implications of the Gradient of Health and Wealth." *Health Affairs* 21(2): 13-30.

Downs, J.S., Lowenstein, G. and Wisdom, J. (2009). "The Psychology of Food Consumption: Strategies for Promoting Healthier Food Choices." *American Economic Review: Papers and Proceedings* 99(2): 159-164.

Elbel, B., Kersh, R., Brescoll, V.L. and Dixon, L.B. (2009). "Calorie Labeling and Food Choices: A First Look at the Effects on Low-Income People in New York City." *Health Affairs* 28(6): w1110-w1121.

Falba, T. and Busch, S. (2005). "Survival Expectations of the Obese: Is Excess Mortality Reflected in Perceptions?" *Obesity Research* 13(4): 754-761.

Fan, M. (2010). "Do Food Stamps Contribute to Obesity in Low-Income Women? Evidence from the National Longitudinal Survey of Youth 1979." *American Journal of Agricultural Economics* 92(4): 1165-1180.

Finkelstein, E.A., Fiebelkorn, I.C. and Wang, G. (2003). "National Medical Spending Attributable to Overweight and Obesity: How Much, and Who's Paying?" *Health Affairs* W3: 219-226.

Finkelstein, E.A., Trogdon, J.G., Cohen, J.W. and Dietz, W. (2009). "Annual Medical Spending Attributable to Obesity: Payer-and Service-Specific Estimates." *Health Affairs* 28(5): w822-w831.

Fox, M.K., Hamilton, W. and Lin, B.-H. (2004). "Effects of Food Assistance and Nutrition Programs on Nutrition and Health: Volume 3 Literature Review." Food Assistance and Nutrition Research Report No. 19-3 Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture.

Fox, M.K., Hamilton, W. and Lin, B.-H. (2004). "Effects of Food Assistance and Nutrition Programs on Nutrition and Health: Volume 4, Executive Summary of the Literature Review." Food Assistance and Nutrition Research Report No. 19-4 Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture.

Freebairn, J. (2010). "Taxation and Obesity?" *The Australian Economic Review* 43(1): 54-62.

Garson, A. and Engelhard, C.L. (2007). "Attacking Obesity: Lessons from Smoking." *Journal of the American College of Cardiology* 49(16): 1673-1675.

Gelbach, J., Klick, J. and Stratmann, T. (2007). "Cheap Donuts and Expensive Broccoli: The Effect of Relative Prices on Obesity." FSU College of Law, Public Law Research Paper 261. Florida State University.

Gibson, D. (2003). "Food Stamp Program Participation is Positively Related to Obesity in Low Income Women." *Journal of Nutrition* 133: 2225-2231.

Grunert, K.G. and Wills, J.M. (2007). "A Review of European Research on Consumer Response to Nutrition Information on Food Labels." *Journal of Public Health* 15: 385-399.

Guthrie, J.F. (2004). "Understanding Fruit and Vegetable Choices: Economic and Behavioral Influences." Agriculture Information Bulletin No. 792. U.S. Department of Agriculture Economic Research Service.

Guthrie, J.F., Frazao, E., Andrews, M. and Smallwood, D. (2007). "Improving Food Choices - Can Food Stamps Do More?" *Amber Waves* 5(2): 22-28.

Institute of Medicine (2002). *The Future of The Public's Health in the 21st Century*. Washington, D.C.: The National Academies Press.

Institute of Medicine and The National Academy of Sciences (1988). *The Future of Public Health*. Washington, DC: National Academy Press.

Institute of Medicine Committee on Food Marketing and the Diets of Children and Youth (2005). "Food Marketing to Children and Youth: Threat or Opportunity?" Washington, D.C., Institute of Medicine.

Interagency Working Group on Food Marketed to Children (2011). "Preliminary Proposed Nutrition Principles to Guide Industry Self-Regulatory Efforts: Request for Comments." CDC, FTC, FDA, and USDA.

Jacobson, M.F. and Brownell, K.D. (2000). "Small Taxes on Soft Drinks and Snack Foods to Promote Health." *American Journal of Public Health* 90(6): 854-857.

Jolliffe, D. (2010). "Overweight and Poor? On the Relationship between Income and Body Mass Index." IZA Discussion Paper No. 5366 Bonn, Germany, Institute for the Study of Labor (IZA).

Just, R.E., Hueth, D.L. and Schmitz, A. (2004). *The Welfare Economics of Public Policy*. Cheltenham, UK: Edward Elgar Publishing Limited.

Kuchler, F., Tegene, A. and Harris, J.M. (2004). "Taxing Snack Foods: Manipulating Diet Quality or Financing Information Programs?" *Review of Agricultural Economics* 27(1): 4-20.

Lin, B.-H. and Guthrie, J.F. (2007). "How Do Low-Income Households Respond to Food Prices." Economic Information Bulletin No. 29-5. U.S. Department of Agriculture Economic Research Service.

Ludwig, D.S. and Pollack, H.A. (2009). "Obesity and the Economy: From Crisis to Opportunity." *Journal of the American Medical Association* 301(5): 533-535.

Mathios, A.D. (2000). "The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market." *Journal of Law and Economics* 43: 651-677.

Meyerhoefer, C.D. and Pylypchuk, Y. (2008). "Does Participation in the Food Stamp Program Increase the Prevalence of Obesity and Health Care Spending?" *American Journal of Agricultural Economics* 90(2): 287-305.

Miao, Z., Beghin, J.C. and Jensen, H.H. (2010). "Taxing Sweets: Sweetener Input Tax or Final Consumption Tax?" Selected paper prepared for presentation at the Agricultural & Applied Economics Association 2010 AAEA, CAES & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010.

Miller, J.C. and Coble, K.H. (2007). "Cheap Food Policy: Fact or Rhetoric?" *Food Policy* 32: 98-111.

Miller, J.C. and Coble, K.H. (2008). "An International Comparison of the Effect of Government Agricultural Support on Food Budget Shares." *Journal of Agricultural and Applied Economics* 40(2): 551-558.

Mojduszka, E.M. and Caswell, J.A. (2000). "A Test of Nutritional Quality Signaling in Food Markets Prior to Implementation of Mandatory Labeling." *American Journal of Agricultural Economics* 82: 298-309.

Muller, M., Schoonover, H. and Wallinga, D. (2007). "Considering the Contribution of U.S. Food and Agricultural Policy to the Obesity Epidemic: Overview and Opportunities." Minneapolis, MN, Institute for Agricultural and Trade Policy.

Nayga, R. (2000). "Schooling, Health Knowledge and Obesity." *Applied Economics* 32: 815-823.

Nestle, M. (2002). *Food Politics*. Berkeley, CA: University of California Press.

Nestle, M. and Jacobson, M.F. (2000). "Halting the Obesity Epidemic: A Public Health Policy Approach." *Public Health Reports* 115: 12-24.

O'Donoghue, T. and Rabin, M. (2006). "Optimal Sin Taxes." *Journal of Public Economics* 90: 1825-1849.

Okrent, A. and Alston, J.M. (2012). "The Effects of Farm Commodity and Retail Food Policies on Obesity and Economic Welfare in the United States." *American Journal of Agricultural Economics,* forthcoming.

Parks, J.C. (2011). "The Effects of the Food Stamp Program on Energy Balance and Obesity." Dissertation, Department of Agricultural and Resource Economics Davis, CA, UC Davis.

Parks, J.C., Alston, J.M. and Okrent, A.M. (2011). "The Marginal Social Cost of Obesity: Measures of the Impact of Changes in Obesity Rates on Public Health-Care Costs." Unpublished Working Paper Davis, CA, Department of Agricultural and Resource Economics, University of California, Davis.

Philipson, T.J. and Posner, R.A. (1999). "The Long-Run Growth in Obesity as a Function of Technological Change." NBER Working Paper Series No. 7423. National Bureau of Economic Research.

Philipson, T.J. and Posner, R.A. (2003). "The Long-Run Growth in Obesity as a Function of Technological Change." *Perspectives in Biology and Medicine* 46(Number 3 Supplement): S87-S107.

Pollan, M. (2003). "The (Agri)Cultural Contradictions of Obesity." *The New York Times*. New York, NY, October 12.

Pollan, M. (2007). "You Are What You Grow." *The New York Times* Magazine. New York, NY, April 22.

Pollan, M. (2008). "Farmer In Chief." *The New York Times*. New York, NY, October 9.

Popkin, B. (2010). *The World is Fat*. New York, NY: Avery.

Ralston, K., Newman, C., Guthrie, J. and Buzby, J. (2008). "The National School Lunch Program: Background, Trends, and Issues." Economic Research Report No. 61. U.S. Department of Agriculture Economic Research Service.

Rashad, I., Grossman, M. and Chou, S.-Y. (2005). "The Super Size of America: An Economic Estimation of Body Mass Index and Obesity in Adults." NBER Working Paper No. 11584. National Bureau of Economic Research.

Rickard, B.J., Okrent, A. and Alston, J.M. (2011). "How Have Agricultural Policies Influenced Caloric Consumption Patterns in the United States?" *Health Economics,* forthcoming.

Roberto, C.A., Bragg, M.A., Livingston, K.A., Harris, J.L., Thompson, J.M., Seamans, M.J. and Brownell, K.D. (2012). "Choosing Front-of-Package Labelling Nutritional Criteria: How Smart were 'Smart Choices'?" *Public Health Nutrition* 16: 1-6.

Sacks, G., Rayner, M. and Swinburn, B. (2009). "Impact of Front-of-Pack 'Traffic-Light' Nutrition Labeling on Consumer Food Purchases in the UK." *Health Promotion International* 24: 344-352.

Sacks, G., Veerman, J.L., Moodie, M. and Swinburn, B. (2011). "'Traffic-Light' Nutrition Labeling and 'Junk-Food' Tax: A Modeled Comparison of Cost-Effectiveness for Obesity Prevention." *International Journal of Obesity* 35: 1001-1009.

Schmidhuber, J. (2004). "The Growing Global Obesity Problem: Some Policy Options to Address It." *Electronic Journal of Agricultural and Development Economics* 12: 272-290.

Schroeter, C., Lusk, J. and Tyner, W. (2008). "Determining the Impact of Food Prices and Income on Body Weight." *Journal of Health Economics* 27(1): 45-68.

Senauer, B. and Gemma, M. (2006). "Why Is the Obesity Rate So Low in Japan and High in the US? Some Possible Economic Explanations." Paper presented at the Conference of the International Association of Agricultural Economists, Gold Coast, Australia, August 12-18, 2006

Smith, T.A., Lin, B. and Lee, J. (2010). "Taxing Caloric Sweetened Beverages: Potential Effects on Beverage Consumption, Calorie Intake, and Obesity." Economic Research Report No. 100. U.S. Department of Agriculture Economic Research Service.

Sumner, D.A. (2005). "Boxed In: Conflicts Between U.S. Farm Policies and WTO Obligations." Trade Policy Analysis No. 1632. Washington, DC, Cato Institute.

Swiburn, B., Gill, T. and Kumanyika, S. (2004). "Obesity Prevention: A Proposed Framework for Translating Evidence Into Action." *Obesity Reviews* 6: 23-33.

Tillotson, J.E. (2004). "America's Obesity: Conflicting Public Policies, Industrial Economic Development and Unintended Human Consequences." *Annual Review of Nutrition* 24: 617-643.

Todd, J.E. and Variyam, J.N. (2008). "The Decline in Consumer Use of Food Nutrition Labels, 1995–2006." ERS Research Report No. 63. Washington DC, U.S. Department of Agriculture Economic Research Service.

Tohill, B.C. (2004). "Dietary Intake of Fruit and Vegetables and Management of Body Weight." Background paper for Joint FAO/WHO Workshop on Fruit and Vegetables for Health, September 1–3, 2004.

Turnock, B.J. (2004). *Public Health: What It Is and How It Works*. Sudbury, Massachusetts: Jones and Bartlett.

Variyam, J.N. (2005). "Nutrition Labeling in the Food-Away-From-Home Sector: An Economic Assessment." Economic Research Report No. 4. Washington, DC, U.S. Department of Agriculture Economic Reseach Service.

Variyam, J.N. and Cawley, J. (2007). "Nutrition Labels and Obesity." NBER Working Paper Series No. 11956. National Bureau of Economic Research.

Ver Ploeg, M., Breneman, V., Farrigan, T., Hamrick, K., Hopkins, D., Kaufman, P., Lin, B.-H., Nord, M., Smith, T., Williams, R., Kinnison, K., Olander, C., Singh, A. and Tuckermanty, E. (2009). "Access to Affordable and Nutritious Food—Measuring and Understanding Food Deserts and Their Consequences: Report to Congress." AP036. Washington, DC, U.S. Department of Agriculture Economic Research Service.

Ver Ploeg, M., Mancino, L. and Lin, B.H. (2007). "Food and Nutrition Assistance Programs and Obesity: 1976-2002." Economic Research Report No. 48. Washington, DC, U.S. Department of Agriculture Economic Research Service.

Visscher, T.L.S. and Seidell, J.C. (2001). "The Public Health Impact of Obesity." *Annual Review of Public Health* 22: 355-375.

Vyth, E.L., Steenhuis, I.H.M., Vlot, J.A., Wulp, A., Hogenes, M.G., Looije, D.H., Brug, J. and Seidell, J.C. (2010). "Actual Use of a Front-of-Pack Nutrition Logo in the Supermarket: Consumers' Motives in Food Choice." *Public Health Nutrition* 13(11): 1882-1889.

Wilde, P.E. (2009). "Self-Regulation and the Response to Concerns About Food and Beverage Marketing to Children in the United States." *Nutrition Reviews* 67(3): 155-166.

Section 2

Addictive Behaviors

10

Alcohol Consumption Among Adolescents in Estonia 1994 – 2010

Kersti Pärna[1], Mariliis Tael[2], Inge Ringmets[1] and Katrin Aasvee[2]
[1]Department of Public Health, University of Tartu
[2]National Institute for Health Development
Estonia

1. Introduction

Estonia is the smallest of the three Baltic countries on the east coast of the Baltic Sea with an area of approximately 45 215 square kilometres and a population of 1.34 million (01.01.2011) (Statistics Estonia, 2011) which was an independent state in 1918–1940. Estonia regained its independence in August 1991 after the collapse of the Soviet Union. This had enormous implications for health and for the political and economic transition during the succeeding years. In 2004, Estonia became a member of the European Union associated with introduction to the common European market and general pressure towards convergence in many policy areas.

The years 1991–1994 represented the period of transition. In terms of economic development, the year 1994 was characterized by significant unemployment (7.6%) and the relatively low gross domestic product (GDP) at current prices (1370 Euros per person) (Statistics Estonia, 2011). In the period of economic growth unemployment was the lowest (4.7% in 2006) and GDP the highest (12 161 Euros per person in 2008). During the following economic downturn unemployment increased to 16.9% and GDP decreased to 10 674 Euros in 2010 (Statistics Estonia, 2011).

It is estimated that alcohol consumption is responsible for about 4% of the total disease burden in the world (WHO, 2007). While regular light to moderate alcohol intake is associated with some reduction in total mortality (Di Castenuovo et al., 2006; Rehm et al., 2009), heavy drinking has been regarded as an important contributor to the high premature mortality rates in central and eastern Europe, particularly in the countries of the former Soviet Union (Chenet et al., 1998; Leinsalu, 2002; Leon et al., 1997; McKee et al., 2000; Popova et al., 2007).

Alcohol consumption per capita (based on legal sale) increased from 5.6 litres of pure alcohol in 1994 to 15.2 in 2007 and thereafter decreased to 12.8 in 2009 in Estonia (Estonian Institute of Economic Research, 2008, 2010). Since 2006, the estimations were made about the quantity of alcohol beverages that tourists consume in Estonia and export from Estonia. Hence, the consumption of alcohol by Estonian inhabitants was 12.6 in 2007 and 10.2 in 2009 (Estonian Institute of Economic Research, 2010). In 1996–2006, alcohol consumption increased especially among men in Estonia. The proportion of men drinking more than

280 g pure alcohol per week increased from 7.4% to 16.2%, and the proportion of women drinking more than 140 g pure alcohol per week increased from 2.7% to 4.7% between 1996 and 2006 (Pärna et al., 2010). In 2006, prevalence rate of non-beverage alcohol consumption was 2.3% among men (Pärna & Leon, 2011). In Estonia alcohol-related mortality was the lowest in 1988–1991 (3.5% of all deaths) and increased to 9.1% in 2002–2005 (Rahu & Pärna, 2009). At the same time, in Estonia age-standardized mortality rates of alcoholic liver cirrhosis increased from 9.7 to 37.5 per 100 000 men and from 2.2 to 16.1 per 100 000 women aged 25–64 in 1996–2006 (Pärna & Rahu, 2010).

Although the vast majority of alcohol-related deaths occur in middle-aged and elderly people, alcohol consumption behaviour is undeniably established in adolescence. Early adolescence is a critical time, where behavioural habits (including alcohol consumption) are developing (Lintonen et al., 2000). Moreover, individuals who begin alcohol consumption at a younger age have an increased risk of becoming regular alcohol drinkers in adulthood. According to schoolchildren's self-estimation, they consume alcohol due to boredom, stress and desire to fit in the group (Milgram, 2001). The use of alcohol may become a means of escaping from situations that youth feel powerless to change (United Nations, 2005). As among adults, alcohol consumption among adolescents is associated with road accidents, suicides, depression, memory problems, fighting, rape and unprotected sexual intercourse which could increase the risk of getting infected with HIV and other sexually transmitted diseases (Williams & Knox, 1987). Alcohol consumption may cause decrease of learning ability, which in turn causes learning difficulties due to which adolescent might fall out of school (Scheier et al., 2000). In addition, the existing evidence about the relationship between adolescent drinking behaviour and family related factors are inconsistent and even contradictory. While some studies have identified a higher risk of excessive adolescent drinking behaviour among lower socio-economic groups (Geckova et al., 2002; Lintonen et al., 2000; Lowry et al., 1996), others have found no or even inverse social gradients in schoolchildren's alcohol consumption (Shucksmith et al., 1997). Living in non-intact families and bad family relationships have been risk factors for alcohol drinking among adolescents (Challier et al., 2000; Shucksmith et al., 1997).

The aim of this chapter is to describe trends in the prevalence of alcohol consumption and episodes of drunkenness and to analyze associations between alcohol consumption, episodes of drunkenness and demographic and family related factors among schoolchildren in Estonia in 1994–2010.

2. Material and methods

2.1 Setting and sampling

This study was based on international survey of health behavior among school-aged children (HBSC survey) conducted among 11-, 13- and 15-year-old schoolchildren in 1993/1994, 1997/1998, 2001/2002, 2005/2006 and 2009/2010 academic year in Estonia. To ensure the clarity of work results, hereafter the end year of the study (1994, 1998, 2002, 2006, 2010) was used while referring to the academic year.

HBSC survey is a World Health Organization collaborative study, in 2010 43 countries participated in the survey. The main objective of HBSC survey is to collect high-quality internationally comparable data on schoolchildren's health behavior, health and well-being in their social context (Roberts et al., 2007). The target groups of the survey are 11-, 13- and 15-year-old adolescents, in Estonia respectively schoolchildren from grades 5, 7 and 9.

Schoolchildren fill an anonymous questionnaire in classroom during one school hour. A person outside from school is present, schoolchildren who miss the class are left out from the survey. The questionnaires are sealed in envelopes in front of schoolchildren to ensure the confidentiality of data (Aasvee et al., 2007; Maser, 2004).

Methods used in this survey are the same in all participating countries and this gives a chance to monitor the changes in schoolchildren health, health behavior and social environment over years in different countries (Aasvee et al., 2007). HBSC survey has been conducted in Estonia in the following periods: February 1994 (King et al., 1996), February–March 1998 (Currie et al., 2000), November–December 2001 (Maser, 2004), February–March 2006 (Currie et al., 2008), February–April 2010.

2.2 Participants

Databases from Statistics Estonia were used to compile the sample. Estonian counties were divided into 12 stratums according to language of instruction and urbanization. Sample sizes were calculated according to the number of children in grades 5, 7 and 9 in 12 stratums (Aasvee et al., 2007).

In HBSC survey it is required that the number of schoolchildren in each age group is approximately 1500 (Roberts et al., 2007). To ensure equal inclusion probabilities the selection was made in two phases. In the first phase schools were selected. Inclusion probabilities were equal to the total number of grades 5, 7 and 9 in the selected school. This means that schools having more classes of these grades had higher probability to get into the sample. In second phase one class from each parallel was randomly chosen. While compiling the age groups, 90% of schoolchildren had to be within ±6 months of the mean age for each age group and remaining 10% no more than 12 months from the mean age. In chosen schools the school board decided whether they wanted to participate in the survey or not. Schoolchildren had the right to refuse filling the questionnaire if they themselves or their parents wanted that way. Over years the non-participation rate has been below 0.5% of the whole sample.

Data files from countries that have participated in the survey have been checked and cleaned in Norwegian Social Science Data Services, where international database was created and is preserved (Aasvee et al., 2007).

This chapter focuses on 13- and 15-year-old adolescents self-assessed alcohol consumption and episodes of drunkenness in 1994–2010.

2.3 Measures

2.3.1 Alcohol consumption

Frequency of alcohol consumption variable was based on different alcoholic drinks (beer, wine, strong alcohol, liqueur, cider, light alcoholic beverages, alcopops) consumption question. Possible answers were 'yes, every day', 'yes, every week', 'yes, every month', 'yes, less than once a month' and 'no, never', 'For the purposes of analysis, schoolchildren were categorised as drinking at least weekly (weekly alcohol consumption) or less often.

2.3.2 Drunkenness

Frequency of drunkenness was assessed by asking whether schoolchildren had ever had so much alcohol that they were really drunk. Possible answers were 'no, never', 'yes, once', 'yes, 2–3 times', 'yes, 4–10 times', and 'yes, more than 10 times'. Responses to this question were grouped into two categories: drunkenness never or only once in life and at least two episodes of drunkenness.

2.3.3 Demographic factors

Sex, age and nationality were used as demographic factors. In accordance with *age groups*, schoolchildren were either 13- or 15-year-old, from grades 7 and 9 respectively. According to *ethnicity* two groups were formed: Estonians and non-Estonians (mainly Russians).

2.3.4 Family related factors

Family structure, family wealth, family affluence scale and family relationships were used as family related factors, that influence health and health behaviour of schoolchildren.

Based on *family structure*, schoolchildren were divided into four groups: 1) two biological parents; 2) one parent; 3) one parent and one step-parent; 4) other combinations (i.e. parent and grandparent, foster home etc.).

Based on *family wealth*, schoolchildren responses were distributed into three groups: 1) bad (very bad and bad); 2) average; 3) good (very good and good).

Family affluence scale (FAS) was calculated on the basis of four items 1) does your family own a car (0, 1, 2 or more), 2) how many times did you travel away on holiday with your family during the past 12 months (0, 1, 2, 3 or more), 3) do you have your own bedroom for yourself (0, 1) and 4) how many computers does your family own (0, 1, 2, 3 or more). A composite FAS score was calculated by summing the responses to these four items ranging from 0 to 9 (Richter et al., 2006). The scores were subsequently recoded into tertiles and respondents were divided into three groups respectively. In 2010 the FAS tertiles were as follows: low (0–4 points); middle (5–6 points); high (7–9 points) (Currie et al., 2008).

Family relationships were evaluated on an 11 item scale, where 0 meant very bad relationships and 10 very good ones. Based on the scale, schoolchildren were divided into three groups: bad relationships (0–3); 2) average relationships (4–6); 3) good relationships (7–10).

2.4 Statistical analysis

Prevalence of alcohol consumption and episodes of drunkenness were calculated separately for 13- and 15-year-old boys and girls. Logistic regression analysis was applied to assess the associations between weekly alcohol consumption, at least two episodes of drunkenness and demographic and family related factors among 13- and 15-year-old schoolchildren. Weekly alcohol consumption (I model) and at least two episodes of drunkenness (II model) were used as binary variables and demographic and family related factors were used as independent variables in logistic regression analysis models. Odds ratios (OR) and corresponding 95% confidence intervals (CI) were computed for both models. OR that equals 1 refers to the base comparison group. OR's of weekly alcohol consumption and at least two episodes of drunkenness were adjusted to all demographic and family related factors in logistic regression analysis.

The present analysis is based on 13- and 15-year-olds (N=12244), 5861 boys and 6383 girls through the study waves (Table 1). Questionnaires, where the correspondents hadn't answered about their age (10 questionnaires) and alcohol consumption (99 questionnaires) were left out from the analysis. Questionnaires that lacked information about episodes of drunkenness were left out from the drunkenness analysis (62 questionnaires). Additional questionnaires that lacked information about questions related to demographic and family related factors were excluded from logistic regression analysis.

Statistical analysis was conducted with Stata 10 (Hills & Stravola, 2007).

	1994		1998		2002		2006		2010	
Age group	N	%	N	%	N	%	N	%	N	%
					Boys					
13-year old	523	49.0	367	59.4	689	52.7	721	47.4	688	51.1
15-year old	545	51.0	251	40.6	619	47.3	799	52.6	659	48.9
Total	1068	100	618	100	1308	100	1520	100	1347	100
					Girls					
13-year old	622	49.8	444	57.1	734	53.1	738	48.5	718	49.3
15-year old	626	49.2	333	42.9	648	46.9	783	51.5	737	50.7
Total	1248	100	777	100	1382	100	1521	100	1455	100
Total	2316	100	1395	100	2690	100	3041	100	2802	100

Table 1. Number of respondents by gender, age and study year, HBSC Survey, Estonia 1994–2010

3. Results

3.1 Demographic and family related characteristics of respondents

In 2010, the number of 13- and 15-year-old adolescents studied was almost equal (Table 2). There was about 77% of Estonian and nearly one fourth non-Estonian schoolchildren. About 63% of adolescents were living with both biological parents, approximately 20% with one parent, 15% with parent and a step-parent, and 2% had some other family structure. Almost half of the schoolchildren evaluated their family wealth to be good, 46% found it was average and about 4% said it was bad. Approximately 38% of schoolchildren had high, one third middle and 27% low FAS. About 77% of adolescents evaluated their family relationships good, 19% average and 3% bad.

Variables	Boys (n=1347)	Girls (n=1455)	Total (n=2802)
	%	%	%
Age			
13-year old	51.1	49.4	50.2
15- year old	48.9	50.7	49.8
Ethnicity			
Estonian	76.8	77.6	77.2
non-Estonian	22.5	22.1	22.3
missing	0.7	0.4	0.5
Family structure			
two parents	65.8	60.6	63.1
one parent	18.6	21.0	19.8
parent and step-parent	13.1	15.8	14.5
other	2.2	2.1	2.1
missing	0.4	0.5	0.4
Family wealth			
good	51.8	47.4	49.5
average	44.0	47.6	45.9
bad	3.1	4.5	3.8
missing	1.1	0.6	0.8
Family affluence scale			
low	24.4	29.5	27.1
middle	32.6	33.4	33.0
high	40.2	35.7	37.9
missing	2.8	1.4	2.0
Family relationships			
good	80.5	73.8	77.0
average	16.6	21.9	19.4
bad	2.4	4.1	3.3
missing	0.5	0.2	0.4

Table 2. Distribution of demographic and family related factors among 13- and 15-year-old schoolchildren by gender, HBSC Survey, Estonia, 2010

3.2 Alcohol consumption

Prevalence rate of alcohol consumption has been quite high among Estonian adolescents over study years. After 2006 the prevalence rate decreased in all age groups.

Among 13-year-old boys the alcohol drinking in 2010 (54.8%) was almost on the same level as in 1994, when it was 52.4% (Table 3). Compared to year 2006 the prevalence rate decreased 22.5%.

Although alcohol consumption rate among 13-year-old girls decreased 10.5% after 2006, it was more than 20% higher in year 2010 than in 1994. Also for the first time, alcohol consumption of girls was higher compared to boys in 2010 (64.3% and 54.8%, respectively).

Alcohol consumption	1994		1998		2002		2006		2010	
	N	%	N	%	N	%	N	%	N	%
					Boys					
Yes	274	52.4	266	72.5	520	75.5	557	77.3	377	54.8
every day	2	0.4	5	1.4	19	2.8	19	2.6	15	2.2
every week	35	6.7	21	5.7	82	11.9	66	9.2	39	5.7
every month	54	10.3	47	12.8	116	16.8	131	18.2	84	12.2
seldom	183	35.0	193	52.6	303	44.0	341	47.3	239	34.7
No	249	47.6	101	27.5	169	24.5	164	22.8	311	45.2
Total	523	100	367	100	689	100	721	100	688	100
					Girls					
Yes	260	41.8	279	62.8	482	65.7	552	74.8	462	64.3
every day	1	0.2	4	0.9	5	0.7	9	1.2	4	0.6
every week	8	1.3	10	2.3	50	6.8	47	6.4	53	7.4
every month	38	6.1	38	8.6	83	11.3	103	14.0	107	14.9
seldom	213	34.2	227	51.1	344	46.9	393	53.3	298	41.5
No	362	58.2	165	37.2	252	34.3	186	25.2	256	35.7
Total	622	100	444	100	734	100	738	100	718	100

Table 3. Distribution of alcohol consumption among 13-year-old schoolchildren by gender, HBSC Survey, Estonia, 1994–2010

Alcohol consumption	1994		1998		2002		2006		2010	
	N	%	N	%	N	%	N	%	N	%
					Boys					
Yes	428	78.5	213	84.9	541	87.4	716	89.6	535	81.2
every day	6	1.1	10	4.0	19	3.1	1	2.4	12	1.8
every week	68	12.5	44	17.5	177	28.6	195	24.4	121	18.4
every month	108	19.8	66	27.0	136	22.0	202	25.3	164	24.9
seldom	246	45.1	93	37.1	209	33.8	300	37.6	238	36.1
No	117	21.5	38	15.1	78	12.6	83	10.4	124	18.8
Total	545	100	251	100	619	100	799	100	659	100
					Girls					
Yes	468	74.8	281	84.4	569	87.8	700	89.4	627	85.1
every day	5	0.8	3	0.9	7	1.1	7	0.9	4	0.5
every week	25	4.0	32	9.6	122	18.8	130	16.6	95	12.9
every month	85	13.6	78	23.4	163	25.2	231	29.5	213	28.9
seldom	353	56.4	168	50.5	277	42.8	332	42.4	315	42.7
No	158	25.2	52	15.6	79	12.2	83	10.6	110	14.9
Total	626	100	333	100	648	100	783	100	737	100

Table 4. Alcohol consumption distribution among 15-year-old schoolchildren by gender, HBSC Survey, Estonia, 1994–2010

Alcohol consumption among 15-year-old boys in 2010 (81.2%) was quite similar compared to 1994 (78.5%) (Table 4). The prevalence rate increased until 2006 and then dropped by 8.4% from 89.6% to 81.2% in 2010. Alcohol consumption among 15-year-old girls increased from 1994 to 2006, when it reached 89.4% and then dropped by 4.3%. For the second time during the study period the prevalence rate of alcohol consumption was higher among girls compared to boys: first in 2002, when difference was 0.4% and then in 2010, when difference was 3.9%.

Prevalence rate of weekly alcohol consumption mostly decreased since 2002, except among 13-year-old girls (Figure 1). The prevalence rate among 13-year-old girls increased from 1.5% in 1994 to 7.9% in 2010. Among 13-year-old boys weekly alcohol consumption prevalence rate decreased from 14.7% in 2002 to 7.9% as among 13-year-old girls in 2010. In 2002 31.7% of 15-year-old boys and 20.0% of the same aged girls consumed alcohol weekly. The prevalence rate decreased from 2002 to 2010, 11.5% among 15-year-old boys and 6.6% among 15-year-old girls.

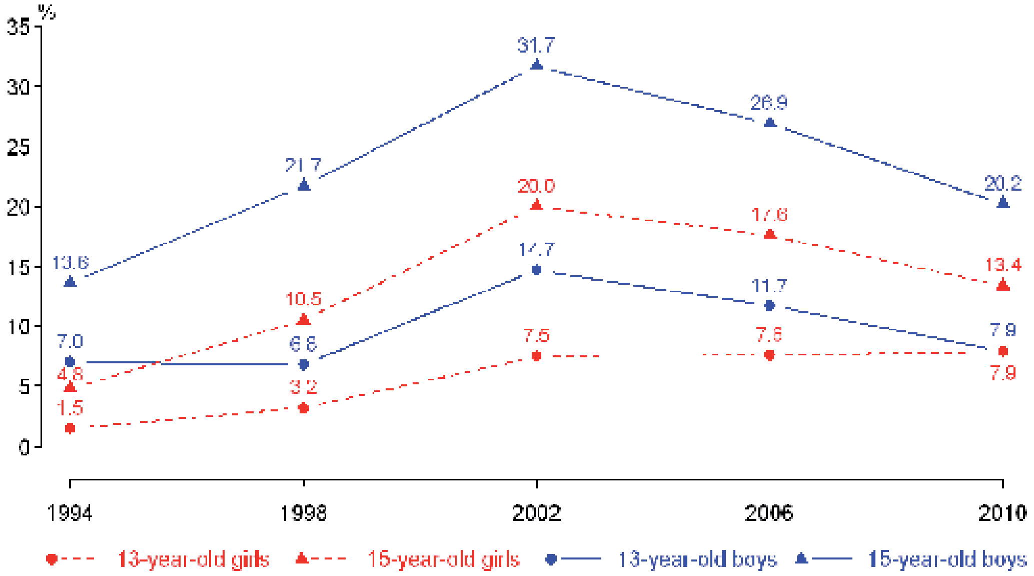


Fig. 1. Prevalence rates of weekly alcohol consumption among 13- and 15-year-old school-children by gender, HBSC Survey, Estonia, 1994–2010

3.3 Drunkenness

The prevalence rate of at least one episode of drunkenness mostly increased from 1994 to 2006 and then decreased. Compared to boys, prevalence rate among girls was only slightly lower in 2010 (Table 5). Among 13-year-old boys the prevalence rate of at least one episode of drunkenness more than doubled in the period of 1994–2006 and after that decreased 11.5% from 45.9% to 34.4%. Among 15-year-old girls the prevalence rate increased since 1994, the difference between years 1994 and 2010 was five-fold (6.9% in 1994 and 31.1% in 2010).

Episodes of drunkenness	1994		1998		2002		2006		2010	
	N	%	N	%	N	%	N	%	N	%
Boys										
Yes	96	18.5	122	33.3	309	44.8	329	45.9	234	34.4
>10 times	5	1.0	3	0.8	37	5.4	40	5.6	28	4.1
4–10 times	6	1.2	11	3.0	33	4.8	43	6.0	29	4.3
2–3 times	24	4.6	41	11.2	100	14.5	102	14.3	66	9.7
once	61	11.8	67	18.3	139	20.2	144	20.1	111	16.3
No	423	81.5	244	66.7	380	55.2	387	54.1	446	65.6
Total	519	100	366	100	689	100	716	100	680	100
Girls										
Yes	43	6.9	66	14.9	209	28.5	219	29.7	221	31.1
>10 times	2	0.3	2	0.5	16	2.2	15	2.0	24	3.4
4–10 times	2	0.3	2	0.5	17	2.3	28	3.8	29	4.1
2–3 times	8	1.3	13	2.9	60	8.2	72	9.8	69	9.7
once	31	5.0	49	11.1	116	15.8	104	14.1	99	13.9
No	577	93.1	377	85.1	525	71.5	519	70.3	490	68.9
Total	620	100	443	100	734	100	738	100	711	100

Tabel 5. Distribution of episodes of drunkenness among 13-year-old schoolchildren by gender, HBSC Survey, Estonia, 1994–2010

Episodes of drunkenness	1994		1998		2002		2006		2010	
	N	%	N	%	N	%	N	%	N	%
Boys										
Yes	256	47.0	150	60.2	460	74.3	566	71.1	445	67.6
>10 times	30	5.5	31	12.5	152	24.6	164	20.6	99	15.1
4–10 times	29	5.3	31	12.5	76	12.3	110	13.8	97	14.7
2–3 times	81	14.9	48	19.3	123	19.9	181	22.7	119	18.1
once	116	21.3	40	16.1	109	17.6	111	13.9	130	19.8
No	289	53.0	99	39.8	159	25.7	230	28.9	213	32.4
Total	545	100	249	100	619	100	796	100	658	100
Girls										
Yes	160	25.6	132	39.6	387	59.8	487	62.4	428	58.2
>10 times	4	0.6	12	3.6	55	8.5	60	7.7	66	9.0
4–10 times	7	1.1	16	4.8	64	9.9	77	9.9	76	10.3
2–3 times	50	8.0	49	14.7	153	23.7	192	24.6	167	22.7
once	99	15.8	55	16.5	115	17.8	158	20.2	119	16.2
No	465	74.4	201	60.4	260	40.2	294	37.6	307	41.8
Total	625	100	333	100	647	100	781	100	735	100

Table 6. Distribution of episodes of drunkenness among 15-year-old schoolchildren by gender, HBSC Survey, Estonia, 1994–2010

Among 15-year-old boys the prevalence rate of at least one episode of drunkenness increased from 1994 (47.0%) to 2002 (74.3%) and then decreased (Table 6). Compared to year 1994, in 2010 the rate was higher by 20.6%. Among 15-year-old girls the prevalence rate increased from 1994 (25.6%) to 2006 (62.4%) and then decreased by 4.2%. The prevalence rate has more than doubled over the whole period.

The prevalence rate of at least two episodes of drunkenness decreased after 2006 among 13- and 15-year-old boys (Figure 2). The prevalence rate among 13-year-old boys in 2006 was 25.8% and in 2010 18.1%. Among 15-year-old boys the rate decreased from 57.2% in 2006 to 47.9% in 2010. Among 13-year-old girls the rate increased since 1994, when it was 1.9%, to 17.2% in 2010. Among 15-year-old girls the rate was stable since 2002, when it was 42.0%. However, compared to year 1994 the rate increased approximately 4 times from 9.8% to 42.0% in 2010.

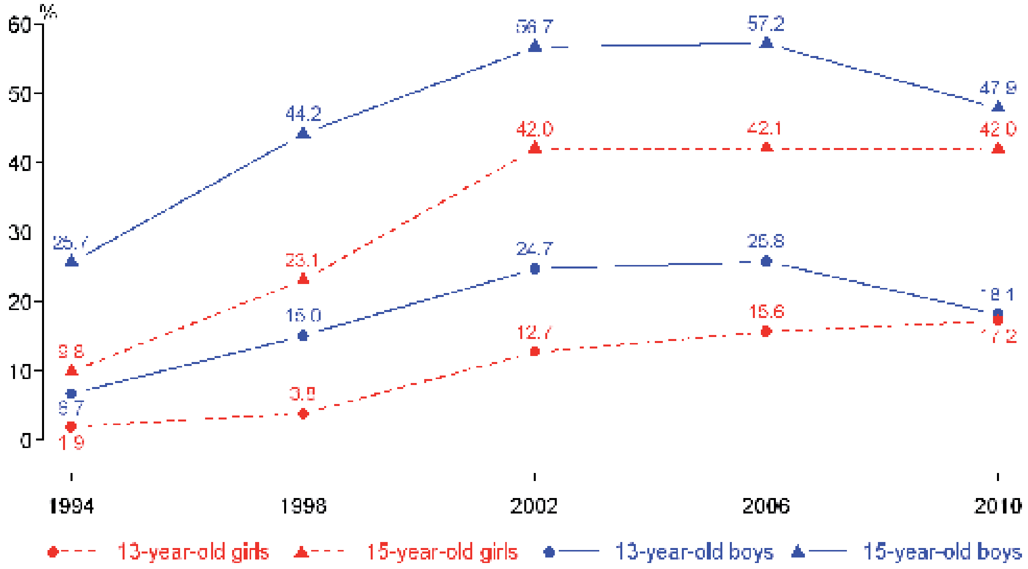


Fig. 2. Prevalence rates of at least two episodes of drunkenness among 13- and 15-year-old schoolchildren by gender, HBSC Survey, Estonia, 1994–2010

3.4 Alcohol consumption and drunkenness by demographic and family related factors

In 2010, weekly alcohol consumption and repeated drunkenness were higher among 15-year-olds (*versus* 13-year-olds) and among Estonian (*versus* non-Estonian) boys and girls (Table 7). Family structure was not associated with drinking alcohol weekly. Compared to schoolchildren living with both parents, odds for repeated drunkenness was higher among girls living with one parent, and among boys and girls living with parent and step-parent. Family wealth and FAS were associated with alcohol consumption and drunkenness only among girls. Odds to drink alcohol weekly and to have at least two episodes of drunkenness was higher among girls living in families with lower perceived wealth, but having higher FAS. Compared to adolescents with the good family relationships, odds to consume alcohol weekly and to have at least two episodes of drunkenness was higher among boys and girls with average and bad family relationships.

Variables	Weekly alcohol consumption (I model)		Repeated drunkenness (II model)	
	Boys	Girls	Boys	Girls
	AOR (95% CI)	AOR (95% CI)	AOR (95% CI)	AOR (95% CI)
Age				
13-year old	1	1	1	1
15- year old	2.78 (1.95–3.96)	1.62 (1.14–2.31)	4.14 (3.18–5.39)	3.47 (2.69–4.48)
Ethnicity				
non-Estonian	1	1	1	1
Estonian	2.63 (1.58–4.36)	2.23 (1.31–3.80)	2.45 (1.74–3.45)	2.03 (1.46–2.83)
Family structure				
two parents	1	1	1	1
one parent	1.21 (0.79–1.88)	1.33 (0.85–2.09)	1.32 (0.95–1.87)	1.55 (1.13–2.13)
parent and step-parent	1.05 (0.64–1.71)	1.32 (0.83–2.09)	1.72 (1.19–2.49)	1.45 (1.04–2.03)
other	1.73 (0.62–4.79)	1.20 (0.39–3.70)	2.01 (0.82–4.90)	2.65 (1.19–5.92)
Family wealth				
good	1	1	1	1
average	0.83 (0.58–1.18)	1.16 (0.80–1.69)	0.89 (0.67–1.16)	1.44 (1.10–1.87)
bad	1.13 (0.46–2.75)	2.38 (1.15–4.92)	1.12 (0.53–2.36)	1.96 (1.07–3.57)
Family affluence scale				
low	1	1	1	1
middle	0.74 (0.46–1.19)	1.35 (0.84–2.17)	1.12 (0.79–1.60)	1.29 (0.93–1.78)
high	1.41 (0.91–2.20)	2.05 (1.27–3.32)	1.35 (0.95–1.92)	1.95 (1.39–2.73)
Family relationships				
good	1	1	1	1
average	1.71 (1.14–2.58)	2.27 (1.54–3.35)	1.76 (1.26–2.46)	1.85 (1.38–2.48)
bad	3.82 (1.65–8.88)	3.30 (1.68–6.47)	3.43 (1.53–7.71)	2.47 (1.36–4.47)

Table 7. Adjusted odds ratios (AOR) and 95% confidence intervals (CI) for weekly alcohol consumption (I model) and for at least two episodes of drunkenness (II model) pending on demographic and family related factors among 13- and 15-year-old schoolchildren by gender, HBSC Survey, Estonia, 2010

4. Discussion

The present chapter focused on alcohol consumption among adolescents in Estonia in 1994–2010.

4.1 Limitations and strength of the survey

Before discussing the results, one has to consider the limitations of the survey. Limitations are mainly related to questionnaire survey and its validity. Some studies have demonstrated that there is tendency to under-report when asking questions on such a sensitive risk taking behaviour like alcohol consumption (Newell et al., 1999). To deal with this potential bias of self-reporting all possible efforts to guarantee anonymity of responses were made. Despite of some limitations, this study provides basic data and several inferences can be drawn.

Strength of this survey is related to the possibility to make cross-national comparisons as standard approach was employed according to the study protocol to use the same questions in each participating country.

4.2 Trends in alcohol consumption and drunkenness

Weekly alcohol consumption increased from 1994 to 2002 and thereafter decreased among boys and girls in Estonia. Only among 13-year-old girls drinking alcohol weekly increased during the whole study period. Nevertheless, compared to the year 1994, prevalence of weekly drinking was much higher in 2010. In 1994–2002, similar increasing trend of alcohol consumption was found among adolescents in neighbouring countries Latvia and Lithuania (Zaborskis et al., 2006).

Similar pattern emerged with drunkenness. Prevalence of at least two episodes of drunkenness increased from 1994 to 2006 and thereafter slightly decreased among boys and girls. Among 13-year-old girls weekly alcohol consumption increased during the whole study period. Throughout the study period prevalence of drunkenness increased more among girls. Again, compared to the year 1994, prevalence of weekly alcohol consumption was much higher in 2010.

At the same time, Simons-Morton et al. (2009) reported world-wide cross-cultural patterns in alcohol consumption. The Northern European countries showed a declining trend, but Eastern European countries (including Estonia) experienced increasing trend in alcohol consumption and drunkenness. The variability of trends by country might reflect alcohol and marketing policy differences by country. Estonia has experienced significant political and economic changes during the last two decades. A lot of effects on alcohol related issues in Estonia could be associated with economic changes of the former Soviet Union as well as with the joining the European Union in 2004.

Gender differences in alcohol consumption among adolescents were pronounced. Throughout the study period, drinking and drunkenness remained higher among boys compared to girls, but the gap between boys and girls declined and girls appeared to be catching up with boys especially among 13-year-olds. Also, gender gap was shrinking in almost all countries participating in HBSC Survey (Simons-Morton et al., 2009). The variability in trends by gender in Estonia could be due to increased effectiveness of contemporary marketing practices or relative ineffectiveness of policies and programs with girls. However, these changes in prevalence could also be due to changes in the social roles of women, allowing girls greater autonomy and wider range of social options (Rahav et al., 2006).

Prevention of alcohol consumption among adolescents is very important because it prevents problems among adults in the future (Hingson et al., 2006). It has been found that the most effective prevention measures are those that are targeted to schoolchildren and to their parents at the same time (Smit et al., 2008; Wu et al., 2003). However, it has to be taken into account that every country needs it's own implementation of a specific program due to cultural differences (Koning et al., 2010). Simons-Morton et al. (2010) compared alcohol consumption of adolescents in USA, Canada and Netherlands and found that there were higher odds to consume alcohol and get drunk in the Netherlands, where minimum age to purchase alcohol is 16 compared to Canada (minimum age 19) and USA (minimum age 21).

4.3 Alcohol consumption and drunkenness by demographic and family related factors

Weekly alcohol consumption and repeated drunkenness were associated with demographic factors like gender and age as described above. In addition, compared to non-Estonian adolescents alcohol drinking was more pronounced among Estonians. At the same time, weekly alcohol consumption was higher, but the amount of pure alcohol consumed per week was lower among adults of ethnic majority in Estonia (Pärna et al., 2010). However, in Lithuania, in schools with Lithuanian teaching language, Polish and Russian adolescents were more exposed to alcohol consumption (Šumskas et al., 2010).

Drunkenness, but not weekly alcohol consumption was associated with family structure among boys and girls. Living in non-intact families was a risk factor for repeated drunkenness. Also, Bjarnason et al. (2003) reported that schoolchildren living with both biological parents engaged less frequently in heavy alcohol consumption than those living in other arrangements. According to world-wide literature, higher supervision in intact families and supportive family environment might be associated with lower alcohol consumption among adolescents in these families (Cookston, 1999; Shucksmith et al., 1997).

Similar association was found between alcohol consumption, repeated drunkenness and family wealth among girls. Lower family wealth was a risk factor for heavy alcohol use among girls in Estonia. Zaborskis et al. (2006) found inverse relationship between alcohol consumption and the perceived family wealth in all three Baltic countries. In 1994–2002, girls in Estonia, but boys in Latvia and Lithuania from the families perceived by them as wealthy were more likely to drink weekly as compared to adolescents from the families perceived by them as not wealthy. This inconsistency in these findings might be explained by time difference in these studies (they were conducted in different years).

Weekly alcohol consumption as well as repeated drunkenness was associated with family affluence scale among girls, but not among boys. At the same time, there was not found any relationship between repeated drunkenness and FAS among 11–15-year-olds in Estonia in 2002 (Richter et al., 2006). Again, this inconsistency in these results might be explained with different age groups and study years used in these studies. Moreover, Richter et al. (2006) found very limited evidence for a close consistent relationship between episodes of drunkenness and parental FAS in almost all other participating countries (27) in the same study. Kuntsche et al. (2004) have pointed out that while adults problem drinking seems to be more common in less affluent groups, this direction might be reversed for adolescents, where accessibility to financial resources is more limited. Also, higher FAS as well as family wealth might be not directly associated with higher education. Inconsistency in association of alcohol drinking with family wealth and FAS could be explained by the fact, that distribution of perceived family wealth among adolescents is quite different as compared to FAS. Adolescents who estimated the family wealth 'not so well-off' or 'not at all well-off' was in 2010 survey 3.2%, by FAS tertiles low economic situation was in 27% of students. Evidently, FAS shows more objectively the financial situation of the family.

In addition, there might be other factors during adolescence, which may have a greater impact on drinking behaviour than family related factors and parental socioeconomic status. For instances, the specific character of adolescence as a stage of experimenting with behaviours associated with adult status (Richter et al., 2006) or situations that youth feel powerless to change (United Nations, 2005).

Weekly alcohol consumption and repeated drunkenness were associated with worse family relationships among boys and girls in this study. According to world-wide literature, a supportive family environment (Shucksmith et al., 1997) and parent-child communication is associated with lowered prevalence of alcohol consumption (Luk et al., 2010).

4.4 Alcohol policy in Estonia

During 1996–2006 alcohol policy was virtually nonexistent in Estonia. There was a national alcoholism and drug abuse prevention programme for 1997–2007, which was continued since 2004 under national drug abuse prevention strategy (1994–2012). This programme mainly focused on the creation of a nationwide information system to evaluate the damage caused by alcohol and drug abuse.

In 1996–2006 the prices of alcoholic beverages increased in Estonia (1.3 times for domestically produced beer and 1.4 times for vodka). However, the average price increase has been slower than the increase in the consumer price index, as well as slower than the income increase of inhabitants, which most likely has also contributed to the increase of alcohol consumption (Estonian Institute of Economic Research, 2008).

Compared to the 2004 level, excise tax increased 45% by the beginning of 2010. The highest tax increases (30% altogether) occurred in 2008 when the economic crisis started to affect the Estonian economy (Lai & Habicht, 2011). This was the first time when affordability of alcoholic beverages decreased after many years.

A nation-wide restriction on the time of off-premise sales of alcoholic beverages was introduced in the summer of 2008. Currently, off-premise sale of alcoholic beverages is prohibited from 10 p.m. to 10 a.m. throughout Estonia.

There is still wide availability of alcohol sales outlets: 198 alcohol retail shops per 100 000 inhabitants in Estonia in 2010, while 6.5 in Finland and 4.5 in Sweden (Estonian Institute of Economic Research, 2010).

Since 2008 advertising of alcoholic beverages on television and radio has been prohibited from 7 a.m. to 9 p.m. in Estonia. Before, prohibition of TV advertisement ended at 8 p.m. for alcoholic beverages other than spirits (Lai & Habicht, 2011).

Interventions on a personal level such as alcohol counselling have had very low focus in Estonia. A pilot study to evaluate the methods of early identification of risk drinking and counselling in the primary health care was carried out in the beginning of 2010. It was found that early identification of risk drinking and counselling are feasible. Continuous practical training, considering the specificity of primary health care, is necessary in order to promote the counselling skills of primary health care specialists (Saame et al., 2011).

In Estonia, further alcohol policy actions should include the reduction of the density of alcohol outlets, more comprehensive advertisement bans, clearer separation of alcoholic beverages from other goods in retail stores and full implementation of brief alcohol interventions in primary health care (Lai & Habicht, 2011).

In addition to previously mentioned policy actions, there should be an alcohol consumption prevention program targeted to adolescents and their parents. Higher price for light alcoholic beverages and consistently asking for ID-card while selling alcohol would lower alcohol consumption among adolescents. Root beer (0.5% and 0.8%) and non-alcoholic cider (1.2%) should be considered as alcoholic beverages. Also raising the minimum age to purchase alcohol from 18 to 21 would be an effective preventive measure.

5. Conclusion

Alcohol consumption among adolescents in Estonia is a serious public health problem. Demographic and family related factors influence alcohol use of adolescents, especially among girls. The results of this study may guide the development of policy and interventions tackling alcohol consumption among adolescents in Estonia.

Additional research is needed on the nature of differences in drinking attitudes and patterns among boys and girls and their parents. Also, it might be interesting to look at the alcohol consumption within the context of other, possibly more relevant factors such as peer or school influence and other parental factors.

6. Acknowledgment

This study was carried out in the framework of the Health Behaviour in School-aged Children (HBSC) Survey, WHO/EURO collaborative international study. This chapter reports data from Estonia (Principal Investigator Katrin Aasvee). The International HBSC Coordinator is Candace Currie (University of St Andrews School of Medicine, St Andrews, Scotland). The Data Management Centre Manager is Oddrun Samdal (University of Bergen, Bergen, Norway). For details, see www.hbsc.org.

This study was supported by the Estonian Ministry of Education and Science (target funding SF0180060s09) and the Estonian Ministry of Social Affairs (Strategy of Prevention of Cardiovascular Diseases 2005–2020).

7. References

Aasvee, K., Poolakese, A., Minossenko, A. & Kurbatova, A. (2007). *Eesti kooliõpilaste tervisekäitumise uuring, 2005/2006. õppeaasta. Tabelid (tulemused kaalutud andmete alusel). [Health behaviour of school-aged children among adolescents in Estonia, 2005/2006. Tables (results based on weighted data)].* Tallinn, Estonia

Bjarnason, T., Andersson, B., Choquet, M., Elekes, Z., Morgan, M. & Rapinett, G. (2003). Alcohol culture, family structure and adolescent alcohol use: multilevel modeling of frequency of heavy drinking among 15–16 year old students in 11 European countries. *Journal of Studies on Alcohol*, Vol.54, No.2, pp. 200–208

Challier, B., Chau, N., Predine, R., Choquet, M. & Legras, B. (2000). Associations of family environment and individual factors with tobacco, alcohol, and illicit drug use in adolescents, *European Journal of Epidemiology*, Vol.16, No.1, pp. 33–42

Chenet, L., Leon, D., McKee, M. & Vassin, S. (1998). Deaths from alcohol and violence in Moscow: socio-economic determinants. *European Journal of Population,* Vol.14, No.1, pp. 19–37

Cookston, J.T. (1999). Effects on adolescent problem behaviors. *Journal of Divorce & Remarriage,* Vol.32, No.1/2, pp. 107–122

Currie, C., Hurrelmann, K., Settertobulte, W., Smith, R. & Todd, J. (Eds). (2000). *Health and health behaviour among young people. Health behaviour in school-aged children: a WHO cross-national study (HBSC) international report. Health policy for children and adolescents. No. 1.* Copenhagen, Denmark

Currie, C., Gabhainn, S.N., Godeau, E., Roberts, C., Smith, R., Currie, D., Picket, W., Richter, M., Morgan, A. & Barnekow, V. (Eds). (2008). *Inequalities in young people's health: international report from the HBSC 2005/06 survey. Health policy for children and adolescents, No. 5.*Copenhagen, Denmark

Di Castelnuovo, A., Costanzo, S., Bagnardi, V., Donati, M.B., Iacoviello, L. & de Gaetano, G. (2006). Alcohol dosing and total mortality in men and women: an updated meta-analysis of 34 prospective studies. *Archives of International Medicine,* Vol.166, No.22, pp. 2437–2445

Estonian Institute of Economic Research. (2008). *Alkoholi turg, tarbimine ja kahjud Eestis. Aastaraamat 2008. [Alcohol market, consumption and harms in Estonia. Yearbook 2008].* Tallinn, Estonia

Estonian Institute of Economic Research. (2010). *Alkoholi turg, tarbimine ja kahjud Eestis. Aastaraamat 2010. [Alcohol market, consumption and harms in Estonia. Yearbook 2010].* Tallinn, Estonia

Geckova, A., van Dijk, J.P., Groothoff, J.W. & Post, D. (2002). Socio-economic differences in health risk behaviour and attitudes towards health risk behaviour among Slovak adolescents. *Sozial und Praventivmedizin,* Vol.47, No.4, pp. 233–239

Hills, M. & Stravola, B. (2007). *A short introduction to Stata for biostatistics: updated to Stata 10,* Timberlake Consultants, ISBN-10 978-0-9557076-1-2, London, United Kingdom

Hingson, R.W., Heeren, T. & Winter, M.R. (2006). Age at drinking onset and alcohol dependence. *Archives of Pediatrics & Adolescent Medicine,* Vol.160, No.7, pp. 739–746

King, A., Wold, B., Tudor-Smith, C. & Harel, Y. (1996). *The health of youth: a cross-national survey. European Series, No.69.* Copenhagen, Denmark

Koning, I.M., van den Eijnden, R.J.J.M., Engels, R.C.M.E., Verdurmen, J.E.E. & Vollebergh, W.A.M. (2010). Why target early adolescents and parents in alcohol prevention? The mediating effects of self-control, rules and attitudes about alcohol use. *Addiction,* Vol.106, No.3, pp. 538–546

Kuntsche, E., Rehm, J. & Gmel, G. (2004). Characteristics of binge drinkers in Europe. *Social Science & Medicine, V*ol.59, No.1, pp. 113–127

Lai, T. & Habicht, J. (2011). Decline in alcohol consumption in Estonia: combined effects of strengthened alcohol policy and economic downturn. *Alcohol and Alcoholism*, Vol.46, No.2, pp. 200–203

Leinsalu, M. (2002). Social variation in self-rated health in Estonia: a cross-sectional study. *Social Science & Medicine,* Vol.55, No.5, pp. 847–861

Lintonen, T., Rimpelä, M., Vikat, A. & Rimpelä, A. (2000). The effect of societal changes on drunkenness trends in early adolescence. *Health Education Research,* Vol.15, No.3, pp. 261–269

Lowry, R., Kann, L., Collins, J.L. & Kolbe, L.J. (1996) The effect of socioeconomic status on chronic disease risk behaviors among US adolescents. *Journal of the American Medical Association*, Vol.276, No.10, pp. 792–797

Luk, J.W., Farhat, T., Iannotti, R.J. & Simons-Morton, B.G. (2010). Parent-child communication and substance use among adolescents: do father and mother communication play a different role for sons and daughters? *Addictive Behaviour*, Vol.35, No.5, pp. 426–431

Maser M. (2004). *Kooliõpilaste tervisekäitumine. 2001/2002. õppeaasta uuring. [Health behaviour of school-aged children, 2001/2002 survey].* Tallinn, Estonia

McKee, M., Pomerleau, J., Robertson, A., Pudule, I., Grinberga, D., Kadziauskiene, K., Abaravicius, A. & Vaask, S. (2000). Alcohol consumption in the Baltic Republics. *Journal of Epidemiology and Community Health,* Vol.54, No.5, pp. 361–366

Milgram, G.G. (2001). Alcohol influences: the role of family and peers. In: *Learning about drinking,* E. Houghton, & A.M. Roche, (Eds.), 85–108, Taylor and Francis, Philadelphia, USA

Newell, S.A., Girgis, A., Sanson-Fisher, R.W. & Savolainen, N.J. (1999). The accuracy of self-reported health behaviors and risk factors relating to cancer and cardiovascular disease in the general population: a critical review. *American Journal of Preventive Medicine,* Vol.17, No.3, pp. 211–229

Popova, S., Rehm, J., Patra, J. & Zatonski, W. (2007). Comparing alcohol consumption in central and eastern Europe to other European countries. *Alcohol and Alcoholism,* Vol.42, No.5, pp. 465–473

Pärna, K. & Leon, D.A. (2011). Surrogate alcohol drinking in Estonia. *Alcoholism: Clinical and Experimental Research,* Vol. 35, No. 8, pp. 1454–1457

Pärna, K. & Rahu, K. (2010). Dramatic increase in alcoholic liver cirrhosis mortality in Estonia in 1992–2008. *Alcohol and Alcoholism,* Vol.45, No.6, pp. 548–551

Pärna, K., Rahu, K., Helakorpi, S. & Tekkel, M. (2010). Alcohol consumption in Estonia and Finland: Finbalt survey 1994–2006. *BMC Public Health,* Vol.10 (May 19), 261.

Rahav, G., Wilsnack, R., Bloomfield, K., Gmel, G. & Kuntsche, S. (2006). The influence of societal level factors on men's and women's alcohol consumption and alcohol problems. *Alcohol and Alcoholism Supplement*, Vol.41, No.1, pp. i47–i55

Rehm, J., Mathers, C., Popova, S., Thavorncharoensap, M., Teerawattananon, Y. & Patra, J. (2009). Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. *Lancet,* Vol.373, No.9682, pp. 2223–2233

Richter, M., Leppin, A. & Gabhainn SN. (2006). The relationship between parental socio-economic status and episodes of drunkenness among adolescents: finding from a cross-national survey. *BMC Public Health,* Vol.6 (November 28), 289

Roberts, C., Currie, C. & Samdal, O. (2007). Measuring the health and health behaviours of adolescents through cross-national survey research: recent developments in the Health Behaviour in Schoolaged Children (HBSC) study. *Journal of Public Health,* Vol.15, No.3, pp. 179–186

Saame, I., Gluškova,N., Viilmann, K. & Kalda, R. (2011). Prooviuuring alkoholi liigtarvitamise varajase avastamise ja nõustamise metoodika hindamiseks Eesti perearstisüsteemis. [Pilot project to evaluate the methods of early identifi cation of risk drinking and counselling in the primary health care system in Estonia]. *Eesti Arst,* Vol.90, No.5, pp. 216–224

Scheier, L.M., Botvin, G.J., Griffin, K.W. & Diaz, T. (2000). Dynamic growth models of self-esteem and adolescent alcohol use. *Journal of Early Adolescence,* Vol.20, No.2, pp. 178–209

Shucksmith, J., Glendinning, A. & Hendry, L. (1997) Adolescent drinking behaviour and the role of family life: a Scottish perspective. *Journal of Adolescence*, Vol.20, No.1, pp. 85–101

Simons-Morton, B.G., Farhat, T., ter Bogt, T.F., Hublet, A., Kuntsche, E., Gabhainn, S.N., Godeau, E. & Kokkevi, A. (2009). Gender specific trends in alcohol use: cross-cultural comparisons from 1998 to 2006 in 24 countries and regions. *International Journal of Public Health*, Vol.54, Suppl2, pp. 199–208

Simons-Morton, B.G., Pickett, W., Boyce, W., ter Bogt, T.F., Vollebergh, W. (2010). Cross-national comparison of adolescent drinking and cannabis use in the United States, Canada, and the Netherlands. *International Journal of Drug Policy,* Vol.21, No.1, pp. 64–69

Smit, E., Verdurmen, J., Monshouwer, K. & Smit, F. (2008). Family interventions and their effect on adolescent alcohol use in general populations; a meta-analysis of randomized controlled trials. *Drug and Alcohol Dependence*, Vol.97, No.3, pp. 195–206

Statistics Estonia. (2011). Available from www.stat.ee (14.06.2011)

Šumskas, L., Lenciasukiene, I. & Zaborskis, A. (2010). Health behaviour inequalities among Lithuanian, Polish and Russian school-aged children in Lithuania. *Central European Journal of Medicine,* Vol.5, No.1, pp. 97–107

Tuinstra, J., Groothoff, J.W., van den Heuvel, W.J. & Post, D. (1998) Socio-economic differences in health risk behavior in adolescence: do they exist? *Social Science & Medicine*, Vol.47, No.1, pp. 67–74

Zaborskis, A., Šumskas, L., Maser, M. & Pudule, I. (2006). Trends in drinking habits among adolescents in the Baltic countries over the period of transition: HBSC survey results, 1993–2002. *BMC Public Health*, Vol.6 (March 15), 67

United Nations. (2005). *World Youth Report 2005: Today and in 2015.* New York, USA

Williams, F.G. & Knox, R. (1987). Alcohol abuse intervention in a university setting. *Journal of American College Health,* Vol.36, No.2, pp. 97–102

World Health Organization. (2007). *WHO Expert Committee on problems related alcohol consumption.WHO technical report series 944. Second report.* World Health Organization, Geneva: Switzerland,. Available from http://www.who.int/substance_abuse/expert_committee_alcohol_trs944.pdf

11

Public Health and Indigenous Australian Gambling: Risky Lifestyle or Harmless Recreation?

Helen Breen, Nerilee Hing and Ashley Gordon
Centre for Gambling Education and Research,
School of Tourism and Hospitality Management, Southern Cross University, Lismore NSW
Australia

1. Introduction

Gambling and particularly gambling-related problems give rise to a complex range of issues. Understanding gambling issues has tended to focus on the individual and their behaviour in the past but is increasingly being recognised as a public health concern (Blaszczynski & Nower, 2007; Shaffer & Korn, 2002). The public health perspective generally considers gambling behavior as a continuum from recreational gambling to low risk and moderate risk gambling and then problem gambling. This perspective focuses attention on the prevention of gambling problems, on minimization of gambling harms and on treatment for those suffering severe gambling problems (Productivity Commission, 2010). A public health view of gambling invites examination of its influences and impacts on populations and communities. Thus the focus of this chapter is on gambling by Indigenous Australians from several communities but belonging to one tribal group in regional Australia.[1] This chapter also draws on a model of gambling derived from the public health literature to analyse risk and protective factors associated with gambling within these communities. The usefulness of public health concepts and models to the study of gambling are demonstrated in this chapter. As well, it makes an empirical contribution to a little studied area.

2. Background

Gambling has long been recorded as a recreational activity in many cultures. Over 300 years ago, regular visits by Macassan fishermen and traders to northern parts of Australia brought card gambling to Indigenous Australians (Breen, 2008). Card gambling remains a widespread and popular form of contemporary recreation in many Indigenous communities, while the expansion of commercial forms of gambling such as poker machines, casinos, lottery-type products, sports betting and wagering has broadened Indigenous participation in gambling (McMillen & Donnelly, 2008). However, there is limited knowledge about Indigenous gambling or gambling-related problems (Belanger,

---

[1] Being aware of the debate around titles used to describe Aboriginal and Torres Strait Islander Australians, in this chapter we use the terms Indigenous Australian and Aboriginal interchangeably depending on source.

2006, 2011). Internationally, various Indigenous populations appear to have a higher prevalence of problem gambling than the general population (Conner & Taggart, 2009; Dyall, 2010; Williams, Stevens & Nixon, 2011). While research in Australia is limited, results suggest higher problem gambling rates among Indigenous Australians compared to non-Indigenous Australians, although prevalence rates have not been rigorously measured (Stevens & Young, 2009).

Nevertheless, Indigenous Australians are considered an at-risk group for gambling problems, given their social and economic disadvantage. Lower socio-economic groups, particularly those who experience poverty, unemployment, welfare dependence, homelessness and low education, usually have higher rates of gambling-related problems than the general population (Shaffer & Korn, 2002; Volberg & Abbott, 1997). Further, gamblers on low incomes suffering from problems with their gambling experience losses that are borne disproportionately (Shaffer & Korn, 2002). In New Zealand, people in lower socio-economic groups, especially Maori and Pacific Island peoples, tend to experience higher rates of problem gambling (Ministry of Health, 2009). A variety of socio-economic factors negatively affect the health and welfare of some Indigenous Australians. Indigenous Australians generally have a lower life expectancy, higher unemployment records, lower education levels, live with poverty and experience higher levels of psychological distress compared to non-Indigenous Australians (Australian Bureau of Statistics (ABS), 2010; Holland, 2011). Livingstone and Adams (2010) maintain that Australian gaming taxation mainly draws from disadvantaged groups of gamblers, adding to their financial hardship and marginalisation. Thus structural deficiencies and inequalities aligned with individual life circumstances may accumulate and impact on gambling issues for Indigenous Australians.

Given that little research has been conducted into gambling by Indigenous Australians, research on identifying and explaining underlying risk factors that contribute to gambling problems appears to be a useful addition to the sparse knowledge in this field. Additionally, research into protective factors, those that assist gamblers in controlling their gambling appears to make a similar contribution, adding balance to this topic. In this research, problem gambling is defined as 'difficulties in limiting money and/or time spent on gambling which leads to adverse consequences for the gambler, others, or for the community' (Neal, Delfabbro & O'Neil, 2005, p. i).

2.1 Theoretical approach

Risk and protective factors associated with gambling can arise from multiple influences. Identifying and analysing risk factors contributing to gambling-related problems form a major component of public health gambling research (Perese, Bellringer & Abbott, 2005). Risk factors, those that exacerbate negative gambling consequences and encourage further gambling (Thomas & Jackson, 2008), are largely unknown for Indigenous gamblers (Breen, Hing & Gordon, 2010). The situation is similar for protective factors, those that protect or reduce gamblers' exposure to harmful gambling consequences (Breen, 2011). Once identified, risk factors can be appropriately targeted for early intervention and prevention even if causal relationships are not established. Additionally, protective factors, those that assist gamblers to make decisions to protect them from harmful outcomes, can potentially inform appropriate public health promotion and education strategies.

The Model of Influences on Gambling Behaviours and Outcomes has been developed specifically to examine influences on the behaviour of gamblers and the consequences of their gambling activities (Thomas & Jackson, 2004). Three important elements of gambling uptake are integral to this model. These elements are the propensity to gamble, the influence of gambling products and services on gambling behaviour and the consequences of gambling behaviour. Each of these three elements has accompanying risk, moderating and protective factors. Leaving aside moderating factors, risk and protective factors may vary according to the propensity to gamble by different populations and by the nature and availability of different forms of gambling. Risk and protective factors associated with gambling outcomes may encourage further gambling for some gamblers but not for others. Thus, designing appropriate public health strategies to address problematic gambling requires a sound understanding of the risk and protective factors associated with gambling by the targeted population. For a depiction of this model, see Figure 1.

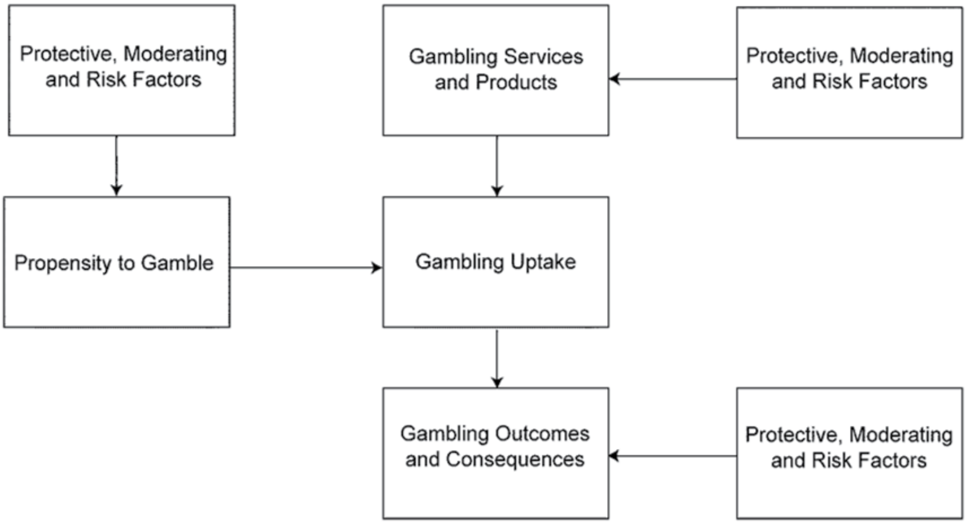


Fig. 1. Model of influences on gambling behaviours and outcomes (Thomas & Jackson, 2004:44)

A variety of influences contribute to gambling uptake and to gambling-related problems. Thomas and Jackson (2004, 2008) propose that behavioural characteristics, sociological and cultural factors strongly influence gambling uptake. In considering gambling uptake and the influence of risk and protective factors associated with it, risks relating to the propensity to gamble might include a faulty understanding about the nature of chance and random events underpinning many forms of gambling, boredom, social isolation, depression and a variety of cultural and ethnic issues (Korn & Shaffer, 1999; Thomas & Jackson, 2004, 2008). Gambling risks relating to the nature of the gambling products and services may include access to gambling, gaming venue characteristics, the nature of the games and the nature of rewards offered (Hing & Haw, 2010; Productivity Commission 1999, 2010). Risk factors potentially affecting gambling outcomes and thus encouraging further gambling might include financial problems, legal troubles, relationship breakdown, depression and/or suicide (Korn & Shaffer, 1999; Thomas & Jackson, 2008; Productivity Commission 1999).

In contrast, protective factors affecting gambling uptake might arise through the propensity to gamble to socialise and relax with others and enjoy any economic outcomes (Korn & Shaffer, 1999; Shaffer & Korn, 2002). For gambling products and services, the influence of responsible gambling strategies publicised by signs and information in the gaming venues, through self-exclusion from venues and from media messages may be protective (QLD Treasury, 2004). For gambling outcomes, having adequate assets to budget for gambling and to provide a buffer against gambling losses (Thomas & Jackson, 2004; Productivity Commission, 1999, 2010) and having family support to help reduce or cease gambling if gambling problems arise (McMillen & Bellew, 2001) may be protective. Any or all of these potential risk and protective factors depend on personal, social, economic and cultural contexts. More risk factors have been identified in prior research than protective factors. In order to provide some balance in this analysis of gambling, the purpose of this chapter is to analyse risk and protective factors associated with gambling by Indigenous Australians in one region of New South Wales (NSW) Australia in order to inform public health measures directed at problem gambling.

3. Methods

An appropriate research design should be culturally sensitive as well as methodologically sound. An interpretative qualitative research design (Guba & Lincoln, 1989) was seen as being culturally appropriate (Martin, 2008). Being based on respectful communication and mutual cooperation, this approach provided rich in-depth data and valuable information. After consultations and discussions, permission was sought and granted by Indigenous Elders in this region and by a university ethics committee for this research to be carried out.

For the setting, in northern New South Wales the eastern corner is comprised of six adjoining local government areas (LGAs). Each LGA has a different geographic, social and economic profile but the Indigenous people of the six LGAs make up one tribal group. To request appropriate permissions for interviews and gather support for the research, we visited the six LGAs several times to explain the project. From an initial list of contacts of Indigenous people and organisations already known, each person interviewed was asked if they could identify other organisations and people relevant to the research. This sampling method, often called snowball sampling, was brokered by someone already known. It was useful in developing confidence and trust between people over time. Snowball sampling was used until all locations were visited and saturation was reached (Creswell, 2007). This process yielded interviews with 169 Indigenous Australians, 21 non-Indigenous Australian gaming venue managers and 10 non-Indigenous Australian gambling counsellors. On average, 30 Indigenous Australians were interviewed in each LGA, although only 20 were interviewed in one LGA where the population was small and widely dispersed. Semi-structured interviews (single or in small groups of two or three) were mainly conducted in workplaces, community and health centres and coffee shops. The three researchers, one Indigenous man and two non-Indigenous women, conducted the one hour interviews in pairs, depending on cultural and gender issues.

All interview notes and recordings were transcribed, then coded with N'Vivo software and analysed using thematic analysis (Braun & Clarke, 2006). The data were coded and analyzed within the larger constructs of the Thomas and Jackson (2004) model for the propensity to gamble, the use of gambling products and services and the consequences of gambling on gambling behaviour. These results are provided below.

4. Results

For the results, risk factors, those that exacerbate and intensify adverse consequences of gambling, are presented first. These are followed by protective factors, those that assist people in making informed choices to protect them from harmful gambling consequences. Please note: all quotations are from Indigenous Australians unless described otherwise.

4.1 Risk factors for indigenous australian gambling

Using the Thomas and Jackson (2004) model, risk factor results were analysed by propensity to gamble, by the use of gambling products and services and by gambling outcomes and consequences.

4.1.1 Propensity to gamble

Risk factors including personal, family, financial, historical and cultural risks were identified by the research participants as increasing the propensity to gamble. Personal risks were seen as gambling to escape from grief and loss, from boredom, due to peer pressure, after consuming alcohol or drugs, from being unemployed and having time to gamble, from having literacy and numeracy problems, and to obtain relief from abuse.

Explaining escape from historical grief and cultural loss, some respondents said ‘memories come up from the past and people push the memories down ... people turn to something, drinking, drugs, gambling, when the old memories surface’. Escape sometimes meant looking for some time away from stressful situations, either at home or at work. Several people mentioned escape in terms of ‘time out alone’ or a break away from heavy responsibilities. Other respondents saw that boredom was a gambling risk especially when people were unemployed, had plenty of free time to gamble and where alternative entertainment options were very limited. Risks arising from peer pressure were explained as ‘(some friends) won’t talk to you unless you’re with them when playing the (poker) machines’. Similar remarks were ‘(I) go to a venue with friends, and the next thing they are all on (poker) machines. If I want to talk to them I have to go over to the (poker) machines too. Then next thing you are putting money in too’.

The use of alcohol and drugs with gambling was described as ‘a cycle involving low self-esteem that leads to drink and drugs’. Some gamblers were seen as being ‘vulnerable to making poor choices, but most want to get out of the cycle’. Several people commented that ‘Aboriginal people are used to being broke and having no money, so when gambling and the money is gone they do not seem to worry’. For some people it was ‘acceptable to be poor in Aboriginal community, where some learnt to live in Struggle Street, used to having nothing’. Explaining budgeting problems and gambling, one person felt many people had ‘not learned to control things in the welfare cycle, they need to learn skills ... need to learn to budget’.

Family risks were said to include generational exposure to gambling, normalisation of youth gambling and a lack of education generally. Gambling was reported to be learnt from parents, family and grandparents where young people learn from exposure and experience. Childhood exposure to gambling and gamblers was apparent to some participants. 'Kids learn the behaviour from parents and older generations ... when it's generational it is very hard to break the cycle'. Some people stated that gambling experiences started early. 'Generations, kids learn to drink and gamble from their parents … then these kids have their own kids and their habits continue'. In one location numeracy and literacy problems were high with one person claiming that not everyone can 'read or write, cannot read signs in the pub (hotel)'. This was further explained as a consequence of some parents not being home to send children to school regularly.

Financial risks were reported as gambling to increase income and to repay debts; and being used to being poor so gambling losses made no great difference to a life of poverty and a cycle of dependency for some. Gambling to increase incomes and reduce financial pressures 'has gone down from generation to generation. It's not dealt with, it's on-going'. Being financially pressed can exacerbate gambling problems leading to 'desperation for money, hoping to win the big jackpot one day'. Some gamblers were reported to believe that a big win would help solve their financial problems, assist them to repay debts and remove the need to borrow from family and friends. Linking gambling to social and economic disadvantage some participants explained 'gambling is linked to unemployment, they have more time to gamble but getting a job is hard'. This can affect self-esteem. 'People think that they are not good enough to get a job, you need experience and no-one gives Indigenous a go'. Summarising financial risks associated with gambling, an Indigenous respondent commented 'Aboriginal people are always broke and are always looking for a quick fix even if they have only $5 left'. A major theme identified by participants was concern about financial risks.

Historical risks were said to include the longevity of gambling as an Indigenous recreation activity and borrowing money for gambling from family and kin. An overwhelming risk was reported thus: 'If you hang out with others at the venue, then you either gamble or are asked for money to support others gambling, especially younger people ... cultural values fall away with alcohol consumption and drugs and gambling'.

Cultural risks were believed to be a loss of Indigenous values, respect and discipline and a deterioration of traditional reciprocity when used for gambling. Cultural losses were seen as diminishing leadership. 'Aboriginal men had a role in society, now they don't ... lost their way'. Others felt there was a 'lack of leadership, there is a struggle to find great leaders in the community, Elders'. Many participants were saddened by the loss of Elders. 'A whole generation of Elders died early, some people have never had Elders to nurture them'. Keeping Indigenous values and culture alive was hard where 'there are less numbers of Elders to get advice'.

Thus, risks associated with propensity to gamble reported by these participants included many personal characteristics, family and cultural factors. More structural risks were linked to financial and historical factors.

4.1.2 Gambling products and services

Risks reported to be associated with gambling products and services were a variety of physical and sensory experiences connected to gambling. These included the sound, light and visual features of poker (slot) machines, their attractive marketing and promotion, and the emotional attachment some people have to a lucky machine or favourite form of gambling. Access to the gambling environment was reported as usually easy and free services enhanced the appeal of staying in a venue to gamble. Social access and acceptance were also risks for some gamblers.

The attraction and risk of using poker machines was explained as 'Aboriginal people like the features of the machines, such as free spins ... like the noise, lights and jackpots'. Others saw attachment to poker machines as being a risk. 'People like to go to the same machines for luck because they think they know what features are needed to win'. Faulty beliefs were evident. 'People believe you win money on machines after a win, that is, have a win and believe it's easy to win (again) so they keep playing'. Superstitious beliefs were apparent when gamblers ask 'What did you come and talk to me for? I was winning; now I have stopped winning because you started talking to me'. As one person explained, 'People don't know the probability of a machine to pay and don't understand the outcomes of losing money'.

Marketing risks were described as 'Clubs always have marketing stuff to keep you in there ... venue advertisements, things you can win, free coffee'. Gambling advertising was seen as 'very attractive and is in your face all the time'. Others reported acceptance as a risk: 'People get sucked in by lights and surroundings of a comfortable venue … air conditioned … no hassles'.

Access to gaming venues, such as hotels or pubs and clubs in urban areas, was generally easy. They provided food, drinks and a variety of hospitality and sports services. Many respondents agreed risks included attractive venue characteristics saying '(Gaming) venues are appealing because you can drink and socialise ... it's all there … meeting place, alcohol, food, gambling' and 'poker machines in the indoor/outdoor area for smokers, very easy to get access'. Therefore, participants reported risks associated with the use of gambling products and services as physical and sensory features linked to gambling as well as freely available physical and social access to gambling.

4.1.3 Consequences of gambling

Risk factors relating to addressing problematic consequences of gambling largely were found to be barriers faced in addressing gambling-related problems. These barriers seemed to fall naturally into two groups, intrinsic or personal barriers and extrinsic or external barriers. Intrinsic barriers were identified as shame, denial, depression and social and cultural norms. In the majority of interviews shame, denial and a subsequent loss of self-esteem were mentioned as barriers. People seemed to 'find it hard to admit … a problem with gambling'. Some gamblers were said to be too ashamed to ask for help and become defensive about their problems. This was partly due to pride and partly because of historical discrimination suffered by Indigenous Australians. 'Ways to resolve problems and arguments were removed in the oppression of our culture; it was not replaced, left with no way to resolve problem'. For other gamblers, 'depression, schizoid and paranoid' co-morbidities arose as barriers to their seeking gambling help. People with depression were seen to have limited ability to cope with pressure from gambling-related problems.

Indigenous gambling was very much seen as a social and cultural norm. 'Gambling is seen as a normal thing to do. For many people it's a routine to be going down to the pub or club'.

These norms were associated with all age groups. 'Gambling is seen as a common activity for young and older Aboriginal people to do'. One person summarised this risk by saying 'Gambling, a lot of people don't think it's a problem because it's a social thing'. Some Indigenous gamblers were said to be 'comfortable living with low incomes'. Further, based on traditional obligations to share with those asking for help, some gamblers relied on 'extended families who take on the role of feeding kids'. This meant that 'the gambler is not challenged' and that gambling problems were extended by family and cultural norms.

Extrinsic barriers were reported by participants to include a lack of knowledge regarding gambling help services, poor access to gambling help, a lack of culturally appropriate gambling help services, a lack of gambling education and awareness, concern about confidentiality of gambling help services, and a lack of knowledge about self-exclusion from gaming venues. Several participants said there was 'no awareness in community of gambling problems ... no programs in place for alcohol, drugs and gambling'. Additionally, other people maintained that many gamblers would not know where to get gambling help because they personally did not know of any gambling help services in their location. Many participants remarked that there were 'no gambling education and awareness programs to let people know about services'.

Logistical problems with isolation and a lack of transport made it even more difficult for some gamblers to access gambling help. For access to gambling help telephone services, a couple of participants noted that many Indigenous people 'don't want to talk about it (gambling) on the phone' and 'would not ring any support for help'.

A lack of 'culturally appropriate gambling help services' was seen as a barrier. A key issue raised was Indigenous people do not like seeing non-Indigenous people for help. 'They don't like going to non-Aboriginal services'. Additionally, they are 'not comfortable with a non-Aboriginal counsellor'. One person felt this was particularly the case for 'the older ones (who) really like Aboriginal services with Aboriginal staff'. Commenting on the lack of services, participants said generally 'there are not enough services (Aboriginal and non-Aboriginal) to help people with gambling problems'. Many respondents noted that 'Mainstream organisations are not addressing Aboriginal issues … are not culturally appropriate ... not capable of addressing Aboriginal gambling issues'.

Regarding confidential services, Indigenous gamblers were concerned about trust. '(We) won't go into non-Indigenous services because confidentiality is a problem'. Gambling help seeking is hindered because Aboriginal people are 'naturally a shy race' said one non-Indigenous gambling counsellor. 'Aboriginal people do not like to talk to people about their problems ... through history they are told to keep their problems to themselves'. Thus, some people feel unable to share their problems resulting in hiding their gambling problems until a crisis occurred.

Thus, intrinsic risks associated with gambling outcomes and consequences were said to be the personal risks facing some gamblers who were experiencing gambling-related problems but were unable to admit this because of their shame, their depression and social and cultural norms encouraging gambling. Extrinsic risks were reported to be the lack of awareness about and availability of gambling help services, a lack of provision of culturally appropriate gambling help services and counsellors, and concern about confidentiality and possible retribution.

4.2 Protective factors for indigenous gambling

Using the Thomas and Jackson (2004) model, protective factor results were analysed in relation to the propensity to gamble, the use of gambling products and services and gambling consequences.

4.2.1 Propensity to gamble

Protective factors identified in relation to the propensity to gamble were grouped into personal, family, financial and cultural themes. Personal protective factors were seen as self-control in controlling gambling activities, self-respect, having a purpose in life and high aspirations, being employed and earning money, learning from experience and religious beliefs.

Several themes emerged from the interviews in relation to personal protective factors. Some of these were linked to having 'high values' including self-respect and respect for others. Having respect and self-control meant having the willpower to 'control ... set limits' for gambling. Similarly, holding high aspirations, including working for a living, was considered a protective factor reducing the propensity to gamble. 'People who work, learn to control their money ... people who do not work hard to earn their money don't value it as much as someone who works for it'. Thus, 'people who work more tend not to be the regular gamblers. They value their money more than others'. Others also commented that people who are aware of themselves and have a purpose in life 'don't want to waste their money in the poker machines'. People had also learnt not to gamble through past experience because 'people see what's destroyed'. Some spoke about their childhood and how they 'don't want that for my kids, gambling and drinking, as seen in my parents' behaviour'. In two locations, a couple of people spoke about how 'the church and religion plays a big role' as a protective factor for gambling. One person found strength through attending church services, saying 'I've been going to church since I was little' and that 'church is strong here'.

Family protective factors were reported to include having strong family influences, family values upheld and positive extended family or kin relationships. Having family around you was also considered a protective factor, mentioned in over half the interviews. This occurred because 'a partner pulls you up' and because 'some families pull one another into line'. Other people talked about the importance of family values. 'Kids, family help people see things that are more important ... different home life and values'. Family responsibilities took precedence for many: 'family and kids, put money towards them first'. One person said 'people learn to spend on important things like kids, food ... they budget'. Another said that 'children change your life as a number one priority'. In addition, the extended family was also identified as being involved 'maybe an aunt or uncle who takes a bit of time' and 'everyone is there to help each other out' and were also seen as protective factors for gambling. Some respondents advocated alternative family activities. 'To do something else, have an open mind, spend time with family, social outings, education and parenting classes'. That is, 'Aboriginal people should look to replace gambling with something else – other options to relax, escape when bored'.

Financial protective factors were said to be education and skills in financial management and small stakes betting. The importance of budgeting skills and money management was discussed in terms of paying bills first before gambling and also setting a gambling budget.

'I set myself a limit of $20 then and still stick to it now' and 'I set a limit, pay bills first and only gamble with what is left over'. Using strategies to control expenditure included 'leaving the key (debit and credit) card at home to limit the money spent' and another was '(to) give his money to his dad who helps him budget'. Other people set up automatic deductions from their welfare pension to ensure essential items were paid first and then spent the remaining amounts as discretionary spending. One person commented 'I just don't like losing money … when you weigh up the odds and how many times people win and how many times people don't, it doesn't add up'. Thus, being able to prioritise spending through effective money management helped people '(to) decide to put money to good use, know how to do this'. Financial education and management was recognised as being important.

Cultural protective factors were believed to be Indigenous Elders acting as role models and people respecting and adhering to traditional Elders' example, cultural cohesion and cultural values. As a source of power and wisdom, 'Elders play an important role as a source of advice and authority'. As role models, Elders were reported as 'leaders and role models, able to correct people who stray off the path'. With good role models 'kids grow up to value hard work and money, education, good food, have a vehicle, employment'. Cultural cohesion was considered to be very important and there was a strong Elders' presence in several communities which was not so evident in others. In locations where Elders authority was respected and heeded, 'cultural values were held by both men and women'. Some people were said to 'come back to [the] community to get help, get advice from Elders … Elders point people in the right direction to get help'. Here, it was said '[we] look after our own'.

4.2.2 Gambling products and services

There were very few protective factors seen as being relating to gambling products and services and those that were mentioned were only suggestions, not really part of people's experience. Some participants mentioned that developing an understanding of the odds of winning would be protective. A few others saw tighter legal controls over gambling spending and reduced attractiveness of poker machines as potentially protective. For example, some participants said that it was protective to 'learn more about how you don't win on a poker machine', although another disagreed, saying that 'knowing the odds wouldn't help much as people still want to escape'. A few people acknowledged that 'legislation really helps with limits on spending and this should be tighter'. Another suggested that 'making (poker) machines less attractive … reduce visual stimuli'. In many ways, these protective factors were potential interventions for reducing gambling risks.

4.2.3 Gambling consequences

Protective factors relating to the consequences of gambling were reported to include factors that facilitated people to be able to address their gambling-related problems. Protective factors were identified as the provision of culturally appropriate gambling help services and counsellors, provision of gambling education and awareness specifically for Indigenous people, encouragement of male role models as community leaders, and a combination of Indigenous community strength and support for people who need assistance. Like gambling products and services, protective factors associated with gambling consequences were mainly seen as being potentially protective because very few had been experienced by these participants.

Most participants raised concerns regarding a lack of Indigenous gambling help services and Indigenous counsellors. They commented 'people did not go to non-Indigenous services before now ... need more Indigenous services'. Some felt it was important to provide a 'non-judgemental service that needs to be in an appropriate place and service'. Then Indigenous gamblers may attend 'if people knew about it and it was culturally appropriate'. In regards to community education about gambling and its impacts, it was contended that the 'Aboriginal population needs to be informed about health promotion. Intervention services could be promoting this on family days and (at) public health events'. Additionally, some suggested 'group awareness sessions would be good, as when it's individual (sessions) they think they are being singled out'. Many participants agreed that Indigenous people 'need education about gambling ... money matters … financial education'. Another way to inform Indigenous gamblers was said to be through 'clubs and pubs … like the smoking campaigns … they could do something similar'.

Where gambling was seen as somewhat problematic in two locations, one facilitator seen to strengthen the community was having strong male role models. Several people said 'men are missing from … this community'. There were 'very few men here as role models'. Men were missing due to young deaths, accidents and incarceration with 'some young boys, teens are in juvenile detention centre'. While very capable women Elders were leading one community, there was a 'need for community leaders (to) build culture' for men to balance the leadership roles and to make up the 'loss of (male) culture'. The need for male role models was seen as one way to strengthen communities and provide young males with aspirations to become Elders.

Another facilitator for addressing gambling problems raised in the interviews was said to be Indigenous community strength and support. Community strength would facilitate group solutions. 'Community groups get together and work together. Get all groups together, talk about issues and target problems early to stop problems'. An example of support was recounted as a gambler who 'tried to kick the habit and came to live here and gave it up'. In this case the influence of a strong community helped this person because they 'dealt with this at a community level'.

Some of these protective factors were potentially helpful interventions for reducing gambling risks. These included the provision of culturally appropriate gambling help services and counsellors and the provision of gambling education and awareness specifically for Indigenous people. More general protective factors such as the encouragement of male role models as community leaders and Indigenous community strength to support for people who need assistance were more general protective factors which could be related to any issues.

4.3 Suggested interventions to address gambling issues

Of the numerous suggestions for potential gambling interventions, the most important and frequently mentioned appeared to be the need for relevant and appropriate community education and awareness programs about gambling, for culturally appropriate gambling help services and trained Indigenous counsellors, and Indigenous specific responsible gambling resources. Indigenous role models were seen as assisting in these processes providing people were trained. Many participants remarked it was important to 'make the community aware and educate them about gambling'. A starting point involved first recognising gambling as being a problem for some Indigenous people and 'start talking about gambling and the problems it can cause'. Some suggestions to facilitate community awareness and education included having 'fun days, barbeques, that's where people talk ... in workshops in the schools'. Other suggestions were 'a big youth forum for the community' and 'use (Aboriginal) Lands Council to spread community awareness about gambling'.

In regards to the provision of culturally appropriate gambling help services, many participants stressed the need for these by saying 'people shy away from non-Indigenous services'; 'some need to speak to Indigenous help'; 'you need more Aboriginal people there, whether it's psychologists or mental health workers'; and, 'we need more Aboriginal counsellors'. A participant asked rhetorically 'why won't our people ring that (free hot line) number, because it's got no Aboriginal staff'. Some people had a fear of being stigmatised. Highlighting the effects of historical discrimination, one person said 'got to be Indigenous way, not colonial way'. Several participants made the point that non-Indigenous workers and counsellors should undergo cultural awareness training and that this 'should be compulsory, Blackfellas access community services, therefore these services need to be able to help, therefore they need awareness training'.

Indigenous specific responsible gambling resources were seen as being 'Aboriginal advertising campaigns for gambling'; 'Aboriginal specific signs and messages in (gaming) venues' and 'an Indigenous gambling help line'. A few participants commented that there were 'not enough Aboriginal workers in hospitality who could raise gambling issues for Aboriginal gamblers'. It was felt that the gaming industry and governments 'can provide much more awareness of gambling that is more focused and appropriate to Aboriginal people'. Governments were expected to take the lead in addressing these issues if they were genuinely serious about reducing harm for Indigenous gamblers.

5. Discussion

Clearly in this chapter, more risk than protective factors have been identified and described as being associated with gambling. Risk and protective factors are now discussed using three fundamental elements of the Thomas and Jackson (2004) model, the propensity to gamble, the use of gambling products and services and the consequences of gambling.

5.1 Risk and protective factors associated with the propensity to gamble

Risks that increased the propensity to gamble included personal, family, financial, historical and cultural risks. Personal risk factors centred on gambling to escape (from grief and loss, violence and abuse), and gambling to alleviate disadvantaged structural conditions (unemployment and a lack of education) and gambling under the influence of alcohol or drugs.

The desire to escape was a risk factor that affected some gamblers more than others. For those removed from their families as children and raised in institutions with no access to their homeland and culture, gambling appeared to provide an escape from harsh experiences and hurtful memories. Indigenous Australian researcher Atkinson (2002) agrees that growing up with long-lasting and traumatic effects of discrimination, marginalisation and disadvantage is reflected in illness, dysfunction and dependency for Indigenous Australians. Experiencing lengthy stress and anxiety, Atkinson (2002:91) noted, 'people begin to feel like losers'. Using gambling to escape appeared as a risk associated with gambling by some gamblers.

Gambling was also used by some Indigenous Australians to try to alleviate their disadvantaged structural conditions such as a lack of education and employment, and by inference, living in poverty. Similarly, Dickerson et al. (1996) found that being younger, single, having a low income, not having full-time work, and having fewer resources to fall back on were multiple risks linked to Indigenous Australian gambling-related problems. Supporting evidence was also reported in New Zealand and Canada regarding gambling by Maori, Pacific and First Nation peoples (Ministry of Health, 2009; Williams et al., 2011) Thus, a lack of education and subsequent unemployment opportunities, underemployment and time to gamble were recognised in the current study as risk factors which increased the propensity for gambling.

Previous research has noted that Indigenous gambling-related risks are heightened when alcohol is a significant problem (Aboriginal and Medical Research Council of New South Wales (AHMRC), 2007). The link between alcohol and drug consumption, Indigenous gambling and negative life events was established by Stevens and Young (2009). A sense of loneliness and internalisation of shame and guilt contribute to some Indigenous people's use of alcohol and gambling in order to reduce their disadvantage (AHMRC, 2007). Self-protective mechanisms can be used by people vulnerable to risk (Dyall, 2010). The co-morbidity between gambling and alcohol use was not unique to this research.

In contrast, personal protective factors appeared to be mainly centred on control, respect and religious beliefs. Similarly, self-control and informal group control protected Indigenous gamblers in north Queensland by reducing their propensity to gamble (Breen 2010). This often included a general agreement on money stakes, on pooled stakes, low denomination gambling and shared winnings. These gamblers regulated their gambling in a healthy, collective way. There was similar evidence of informal group control and self-management found in this current research.

Religious beliefs were found to help people make decisions to abstain from gambling or to gamble in a controlled way. Some people had experienced adverse consequences from earlier gambling and their religious faith helped them resist repeating those experiences. In this research, as in north Queensland (Breen 2010), religious beliefs appeared to provide non-gamblers and reformed gamblers with protection through their beliefs, values and involvement in religious activities.

Family risk factors said to be associated gambling focused on generational issues and normalisation of youth gambling. Family and adult gambling has previously been seen as a model for gambling by younger Indigenous Australians (AHMRC, 2007; McDonald & Wombo, 2006). A family history of gambling is important for linking social relationships with gambling. Maori mothers and grandmothers were said to be role models for their children, socialising them into gambling (Morrison, 1999). Indigenous youth appeared to follow social norms for gambling in family circles. Thus, unhealthy gambling activity by adult gamblers provides an example to others, especially youth, which may reinforce this risky behaviour.

However, protective factors were identified as strong family values and positive relationships. In north Queensland, Breen (2010) found role models to be parents, grandparents, extended family members and siblings. They assisted some gamblers to manage their gambling by advice and substitution of activities, such as fishing. Role models and strong family ties also helped shield some Indigenous people from pressure to gamble.

Some gamblers were said to face financial risks and create a cycle of dependency in their efforts to increase income and alleviate poverty through gambling. This was exacerbated by having few skills in budgeting and financial management. Prior research has also found that Indigenous gamblers on low incomes are more likely to experience problems with their gambling because they can least afford to lose money (AHMRC, 2007). Stevens and Young (2009) found that Indigenous gambling problems were linked to low individual and household income, uncertain housing tenure and cash flow problems. Using gambling to try to increase resources is an unreliable and high risk strategy which may result in gambling with borrowed money. Gambling to make money was a risk associated with gambling.

Conversely, a financial protective factor used by some gamblers in the current study was limiting gambling activities to small stakes betting and having some understanding of the odds of winning. Such limits generally protect against problem gambling, that is, gambling characterised by difficulties in limiting money and/or time spent on gambling which leads to adverse consequences for the gambler, others and the community (Neal et. al. 2005:125). A further protective factor was identified as education and skills in financial management, with some gamblers reporting that they pay their bills first and only gamble with discretionary funds.

Historical and cultural risks were reported to be centred on the longevity of Indigenous activities (cultural acceptance) and a loss of traditional values and respect. Similarly, some gamblers in north Queensland also found it difficult or nearly impossible to avoid gambling activities by virtue of a history of gambling within their families and kin folk (Breen, 2010). When a cycle of gambling losses resulted in debts and loans incurred by chasing losses, the ripple effect of gambling losses was felt by others as gamblers borrowed money, food and other essentials. Reciprocity traditionally supported others in times of genuine need, but was found to be a risk for some gamblers when used to extend their gambling.

However, protective factors included respect for cultural values and Elders' authority. Engagement with culture and spirituality, Atkinson (2002) maintained, are the foundations of health and healing for those traumatised Indigenous Australians seeking escape through highly addictive behaviours. Indigenous role models, particularly the example of Elders in certain communities, were protective by providing cultural leadership when acting in ways that were positive, balanced and respectful of Indigenous values.

5.2 Risk and protective factors associated with the use of gambling products and services

Risks associated with gambling products and services were reported by participants as including the physical and sensory experiences tied to gambling, use of 'lucky' poker machines, marketing and the gambling environment.

Poker machine appeals include fast games, recognisable artwork and graphics, enjoyable sounds, pleasing music, free spins, intermittent payouts, cash prizes and jackpot prizes (Productivity Commission 2010). The participants' mentioned a variety of attractive features making poker machine gambling a very popular form of commercial gambling here. Poker machine gambling has been linked to a heightened risk of developing gambling problems generally and particularly for Indigenous women gamblers (AHMRC 2007). Some gamblers here, both male and female, appeared to be captured by poker machines and were reluctant to stop gambling when other gamblers are unaffected.

A favourable view of gambling was found to be supported by marketing and venue appeal. In research in the Northern Territory, increased Indigenous participation in commercial gambling has been attributed to the socially inclusive nature of the gaming venues (McDonald & Wombo 2006). Similarly, the use of Maori cultural symbols and traditions for marketing gambling businesses, Dyall (2009) argued, encourages Maori gambling. Some marketing and promotional activities were perceived as risk factors by participants in this research.

Easy physical and social access was seen to increase the appeal of gambling, especially if life at home was tense or over-crowded. Gaming venues with comfortable facilities made it easy to stay in a venue and gamble, a finding also noted by the AHMRC (2007). Similarly, in New Zealand, Morrison (2004) found that Maori women sought glamour and comfort in gaming venues as well as an escape from stress. While few protective factors were found associated with the use of gambling products and services, our risk factor results are confirmed by the literature for Indigenous gamblers here.

5.3 Risk and protective factors associated with the consequences of gambling

As noted earlier, risk factors relating to the consequences of gambling were mainly expressed here in terms of barriers to addressing a gambling problem, while protective factors related to facilitators to addressing a gambling problem. A most important intrinsic barrier centred on shame. Similarly, an important extrinsic barrier focused on a lack of culturally appropriate gambling help services.

Shame experienced with gambling-related problems was recognised as a source of pain for Indigenous gamblers in the past (Breen, 2010; McDonald & Wombo, 2006). The main barriers for Indigenous people in help-seeking for gambling problems were shame and some unwillingness to trust the confidentiality of counselling services (AHMRC, 2007). In contrast, Dickerson et al. (1996) reported that, in New South Wales, Indigenous gamblers with gambling-related problems sought help for their gambling at five times the rate of non-Indigenous gamblers with similar problems. However, the type of help sought was not identified. Denial and concealment were reported here as preventing Indigenous gamblers from seeking help and a lack of appropriate gambling help services exacerbated this risk.

In contrast, facilitators suggested by participants were personal recognition of a gambling-related problem. This depends on having a sufficient understanding of gambling and its related effects to recognise signs of a gambling problem and knowing where to seek help. Mainstream research has also shown that gamblers with family and social support have heightened capacity to address gambling-related problems (Thomas & Jackson, 2004). For Indigenous gamblers, supportive attitudes and assurance of family help have also been identified as important facilitators for help-seeking (Breen, 2010).

A lack of culturally appropriate gambling help services was an important extrinsic barrier to seeking help for a gambling problem. Uncertainty was raised about culturally unfamiliar operational processes of the gambling help services, a variety of communication styles, unequal power relationships and the gender of the counsellors. These barriers to help-seeking have been raised previously (McMillen & Bellew, 2001). Facilitators to reduce these barriers included the provision of a range of gambling help services, including culturally appropriate community education and awareness, gambling counselling, trained Indigenous counsellors and for a few, rehabilitation. To reduce gambling risks and improve protection, the location of gambling help services within Indigenous health services could assist gamblers to access help without feeling shame.

Although this research was undertaken in one area of New South Wales and does not represent the state population of Indigenous Australians, it does reveal information on risk and protective factors associated with gambling that has not previously been identified or documented. The qualitative interviews conducted here have brought to light numerous risk and protective factors that deserve further investigation, possibly using quantitative methods to allow for generalisability.

6. Conclusions

The application of a public health model to investigate gambling by Indigenous Australians in one area, northern New South Wales, has highlighted the complexities of Indigenous gambling motivations, behaviours, consequences, risk factors and protective factors. Importantly, this research has pointed out the opportunity for potential interventions to develop culturally sensitive and inclusive responsible gambling strategies and practices reported to be appropriate for Indigenous Australians here. These strategies and practices, developed in cooperation with Indigenous Australians, could filter down into other communities, heightening protection of Indigenous gamblers through the active participation of Indigenous collaborators. It is hoped that this research has provided a useful platform from which such actions can proceed.

7. Acknowledgement

This project and subsequent authorship of this article was funded by Gambling Research Australia.

We acknowledge and thank the Indigenous Australians involved in this research for their generous cooperation and consideration.

8. References

Aboriginal Health & Medical Research Council of New South Wales (AHMRC). (2007). *Pressing problems, gambling issues and responses for NSW Aboriginal communities*, AHMRC, Sydney.

Australian Bureau of Statistics. (ABS). (2010). *The Health and Welfare of Australia's Aboriginal and Torres Strait Islander Peoples*, No. 4704.0. ABS, Canberra.

Belanger, Y. (2006). *Gambling with the Future, The evolution of Aboriginal Gaming in Canada, Purich* Pubishing Ltd., Saskatoon.

Belanger, Y. (2011). (Ed.). *First Nations gambling in Canada: Current trends and issues*, University of Manitoba Press, Winnipeg.

Blaszczynski, A. & Nower, L. (2007). Research and measurement issues in gambling studies: Etiological Models. In G. Smith, D. Hodgins & R. Williams (Eds.). *Research and* *measurement issues in gambling studies*, pp. 323-344, Academic Press, Burlington, MA.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology, *Qualitative research in psychology*, vol. 3, pp. 77-101.

Breen, H. (2008). Visitors to Northern Australia: Debating the history of Indigenous gambling. *International Gambling Studies,* Vol.8, No.2, pp. 137-150.

Breen. H. (2010). *Risk and protective factors associated with Indigenous gambling in north Queensland.* Unpublished PhD thesis, Southern Cross University, Lismore.

Breen, H. (2011). Risk and protective factors associated with gambling consequences for Indigenous Australians in north Queensland. *International Journal of Mental Health and Addiction.* DOI: 10.1007/s11469-011-9315-8 Published Online First 26.2.2011, Retrieved from http://www.springerlink.com/content/l152j81274544p52/ (17 October 2011)

Breen, H., Hing, N. & Gordon, A. (2010). *Exploring Indigenous gambling: Understanding Indigenous gambling behaviour, consequences, risk factors and potential interventions.* Gambling Research Australia, Melbourne.

Conner, T. & Taggart, W. (2009). The impact of gaming on the Indian Nations in New Mexico, *Social Science Quarterly,* Vol.90, No.1, pp. 50-70.

Creswell, J. (2007). *Research design: Qualitative, quantitative and mixed methods approaches* (2nd edition), Sage Publications, Thousand Oaks, CA.

Dickerson, M., Allcock, C., Blaszczynski, A., Nicholls, B., Williams, J. & Maddern, R. (1996). *A preliminary exploration of the positive and negative impacts of gambling and wagering on Aboriginal people in NSW,* AIGR, Sydney.

Dyall, L. (2007). Gambling, social disorganisation and deprivation. *International Journal of Mental Health and Addiction*, Vol.5, No.4, pp. 320-330.

Dyall, L. (2010). Gambling: A poison chalice for Indigenous Peoples. *International Journal of Mental Health and Addiction,* Vol.8, pp. 205–213.

Guba, E. & Lincoln, Y. (1989). *Fourth Generation Evaluation.* Sage Publications, Newbury Park, CA.

Hing, N., Breen, H. & Gordon, A. (2010). Respecting Cultural Values: Conducting a Gambling Survey in an Australian Indigenous Community. *Australian and New Zealand Journal of Public Health,* Vol.34, No.6, pp. 547-553. (DOI: 10.1111/j.1753-6405.2010.00624.x)

Hing, N. & Haw, J. (2010). *Influence of venue characteristics on a player's decision to attend a gambling venue.* Gambling Research Australia, Melbourne.

Holland, C. (2011). *Shadow Report, On Australian governments' progress towards closing the gap in life expectancy between Indigenous and non-Indigenous Australians*. Close the Gap Campaign Steering Committee, Canberra.

Korn, D. & Shaffer, H. (1999). Gambling and the health of the public: Adopting a public health perspective. *Journal of Gambling Studies*, Vol.15, No.4, pp. 289-365.

Livingstone, C. & Adams, P. (2010). Harm promotion: Observations on the symbiosis between government and private industries in Australasia for the development of highly accessible gambling markets. *Addiction,* Vol.106, pp. 3-8.

Martin, K. (2008). *Please knock before you enter, Aboriginal regulation of Outsiders and the implications for researchers*. Post Pressed, Brisbane.

McDonald, H. & Wombo, B. (2006). *Indigenous gambling scoping study - Draft report.* School for Social and Policy Research, Charles Darwin University, Darwin.

McMillen, J. & Bellew, N. (2001). *ACT needs analysis, gambling support services.* AIGR, Sydney.

McMillen, J. & Donnelly, K. (2008). Gambling in Australian indigenous communities: The state of play. *Australian Journal of Social Issues,* Vol.43, No.3, pp. 397-426.

Ministry of Health. (2009). *A Focus on Problem Gambling: Results of the 2006/07 New Zealand Health Survey*. Ministry of Health, Wellington, New Zealand.

Morrison, L. (1999). T*he good and the bad times: Maori women's experiences of gambling in Rotorua*. Unpublished Master's thesis, University of Waikato, New Zealand.

Morrison, L. 2004, 'Pokie gambling and Maori women: Friend or foe?' *eGambling: The Electronic Journal of Gambling Issues,* Vol. 12, Retrieved from http://www.camh.net/egambling/issue12/jgi_12_morrison.html (15 October 2011)

Neal, P., Delfabbro, P. & O'Neil, M. (2005). *Problem gambling and harm: Towards a national definition*, Gambling Research Australian, Melbourne.

Perese, L., Bellringer, M. & Abbott, M. (2005). *Literature review to inform social marketing and approaches, and behaviour change indicators, to prevent and minimise gambling harm,* Health Sponsorship Council, New Zealand.

Productivity Commission. (1999). *Australia's gambling industries*. Inquiry Report No. 10, Ausinfo, Canberra.

Productivity Commission. (2010). *Gambling*. Report no. 50, Canberra.

Shaffer, H., & Korn, D. (2002). Gambling and related mental disorders: A public health analysis. *Annual Review of Public Health*, Vol.*23*, pp. 171-212.

Stevens, M. & Young, M. (2009). *Reported Gambling Problems in the Indigenous and Total Australian Population.* Gambling Research Australia, Melbourne.

Thomas, S. & Jackson, A. (2004). Influences on gambling behaviors and outcomes: A model for the design of effective interventions. *Gambling Research,* Vol.6, No.2, pp. 40-51.

Thomas, S. & Jackson, A. (2008) *Risk and protective factors, depression and comorbidities in problem gambling,* Report prepared for beyondblue, The Problem Gambling Research and Treatment Centre, Melbourne.

Volberg, R., & Abbott, M. (1997). Gambling and problem gambling among indigenous peoples. *Journal of Substance Use and Misuse,* Vol.32, No.11, pp. 1525-1538.

Williams, R., Stevens, R., & Nixon, G. (2011).Gambling and problem gambling in North American Aboriginal People. In Y. Belanger (Ed). *First Nations gambling in Canada: Current trends and issues*, pp. 166-194, University of Manitoba Press, Winnipeg.

12

Self Medication, Drug Dependency and Self-Managed Health Care – A Review

A. O. Afolabi
Department of Dental Services, Federal Medical Centre, Owo, Ondo State
Nigeria

1. Introduction

Craving for medicine and self medication has been part of mankind from one generation to another. People generally hold the view that medicines should be used in the event of any sickness or discomfort[1]. Consumers are being called upon to assume more responsibility for their health promotion and disease prevention practices. This challenge has motivated them to embrace the concept of self medication.

It is a common knowledge that there are not enough Doctors and Pharmacists in Africa and other developing countries to direct and guide everyone who become ill on the correct use of medications. Drug manufacturers have not helped matters as their chief concern is to promote the sale of their medicines without giving adequate information to the public on such drug if possible in the local language. This is compounded by high illiteracy level, poverty and inadequate health facilities and personnel. Self medication offers a way out as people begin to sense the positive benefits of multiplying their options in health care. In the developed countries with sufficient health manpower, many people still buy non-dangerous medications without a doctor's prescription[2,3.] These are the over-the-counter (OTC) drugs whose sales statistics reflect the pattern of self medication[4].

Studies in Britain and United States show that on the average 50-75% of health care takes place within the realm of self medication[5,6]. This practice cuts across culture, gender, health and social status, race, occupation or any other sociodemographic or sociomedical state.

A cost benefit study in the UK concluded that availability of OTC drugs to the public results in saving the General Practitioner's time besides other benefits to the consumer since he /she can attend to other matters at the same time[7]. The total health expenditure within a country's gross domestic product (GDP) strains the public purse to which increasing demand is made. One potential means of reducing this pressure within the health budget is a greater reliance on self health care. Self care users may visit the physician less often and stay fewer days in the hospital resulting in lower expenditure for the hospital and Physician services[8.]

Self medication is however of public health concern because of the problem of drug misuse and abuse and its attendant medical (drug resistance and hypersensitivity), social (juvenile delinquency) and psychological (addiction and physical dependence) problems. In addition, lack of knowledge of possible side effects of self-administered medication and possibility of selling potentially dangerous drugs as over-the-counter in developing countries could have a deleterious effect on the general health of the public.

This paper attempts to review existing information in the literature on the scope and distribution of self medication, its relationship with drug dependency and possible factors which might affect it. Recommendations are made on how self-medication can be effectively utilized in self managed health care.

2. Historical perspective

Man has used drugs for various purposes from the dawn of history. Herbal and other plant derived remedies have been estimated by the World Health Organization (WHO) to be the most frequently used therapies worldwide. Plant-derived remedies can contain chemicals with potent pharmacologic and toxicologic properties[9]. From the ancient civilization of South America came cocaine obtained from the leaves of Erythroxylon coca which was chewed for pleasure and reduction of fatigue. Extracts of cacti and mushroom species, used for religious purposes among Central and North American Indians can be used as an hallucinogenic agent. In Africa, eserine, a component of miotic eyedrops develops from Calabar beans used in fetish practices. Bronchiodilatory effects of ephedra develops from ephedra plants species in ancients China while digitalis, a potent heart stimulant was developed from purple foxglove, an ingredient of herbal folk medicine in England[10.]

Self medication had also been derived from other sources outside plants. In the southern United States of America, certain foods are used to reduce the excess volume of 'blood' which was believed to cause the illnesses; in Latin America, certain foods are used to counteract 'hot' or 'cold' illness and to restore the body equilibrium [11]; in the majority of Xhosa speaking women of South Africa, indigenous healing practices are used for themselves and their babies because of the need to 'strengthen' the womb against sorcery, prevent childhood illness and to treat symptoms they perceive biomedical services would not be able to treat[12].

3. Scope and distribution

The concept of self medication, encouraging an individual to look after minor ailments with simple but effective remedies, has been adopted the world over. People hold the view that medicines should be used in the event of any sickness or discomfort[12]. In the United Kingdom, the government encourages self reliance while agencies like WHO promote individual family and community participation in primary health care[13]. Poor diagnostic ability compounded by a limited knowledge of appropriate management results in the increase of self medication and low rate of health care utilization[14]. People are more likely to seek care from Physicians for symptoms that are serious since it was perceived that Doctors do not have time for trivial complaints[15]. Hence, whenever they perceive a symptom as minor, self medication was usually used for treating themselves[16]. A survey conducted in Poland revealed that self medication, while widespread, does not imply a negative attitude towards health professionals or the existing system of medical care[17]. On the contrary, people began to sense the positive benefits of self care among which is its apparent contribution toward improvement of the efficiency of the over all health care system of themselves[18].

The basic knowledge about the proper way of dealing with drugs and potential dangers of self medication is both insufficient and under estimated. This can be seen from earlier[19] and later[20] studies conducted in Switzerland and Nigeria respectively which revealed that nearly one third of the population lacked sufficient drug knowledge. Lam and co-workers also showed that lack of knowledge was common with its side effects[21].

In spite of the above shortcomings in knowledge, individual attitude to self medication has not diminished as can be seen from various studies conducted worldwide which revealed prevalence which range from 60-90%[22,23]. For instance, Afolabi in a study of market women in a suburban community of Lagos, Nigeria reported 95-98%[24]; Omolase et al established that 79% of ophthalmic patients[25] and 85% of patients in the general out patients clinic[26] in Owo, Nigeria admitted self-medication and Servidoni et al in an Ear, Nose and Throat clinic in Brazil reported 83%[27]. Bamgboye et al, in a study of workers in a tertiary hospital in Nigeria reported a prevalence of 73%[23], Onajole et al[28] established in Lagos, Nigeria that 71% of their respondents admitted to drug misuse, Agbor and his co-worker[29] reported 67.8% prevalence for oral health problems in Cameroun while three studies of different population groups in Sudan reported that 81.8%[30], 79.5%[31] and 73.9%[22] respectively engaged in self medication without prescription or medical advice. However, other studies revealed a much lower prevalence for self-medication. For instance, it was 22% in a population-based study in Czechoslovakia[32], 42% among dental outpatients in Nigeria[33], 32.5% in a study among Hong Kong Chinese population[21], 27.5% in a study conducted among Ethiopian populations[34], 22% among ophthalmic patients in Ibadan, Nigeria[35] and 31% of ear, nose and throat outpatients in Nigeria[36]. The extreme variation in figures might be due to the composition of the sample population, survey location and methods[22].

Majority of those who self medicated reported improvement of their symptoms and this could have accounted for the delay in presentation at the clinic/hospital[22,23]. This was confirmed by a Nigerian study of infants with acute respiratory tract infection which revealed that 32% had been treated with cough medicines, 42% with antipyretics, 5% with antibiotics and 10% with haematinics before they were brought to the clinic[37].

For chronic health problems, people device strategies of self care over months and years and apply them during flare-ups[38]. For instance, a study of asthmatics showed that while 80% of sufferers tended to reduce doses following improvement, 48% of these bought their drugs without prescription for prophylaxis and in case of flare-ups[39]. Among migraine sufferers, 42% self-treated themselves instead of consulting Physicians as most sufferers have learned to live with their condition[40]. This was further confirmed in a study of Canadian migraine sufferers where about 90% used OTC drugs to self-treat their ailments whenever they had an attack[41].

Pharmaceuticals can be bought without a Doctor's prescription for self-treatment in most pharmaceutical shops in developing countries. It was 51% of drug sales in an Ecuadorian study [42], 66.3% from a study in the Phillipines [43] and 80% of drug purchases in a study across the U.S-Mexico border[3.] In countries where drug purchase is regulated like Portugal, a reduced prevalence of 26.2% was reported[44].This emphasizes the importance of careful drug history for General Practitioners and Physicians so as to be aware of what patients are taking before treatment commence especially where subtherapeutic doses are involved. For instance, Bosch and co-workers reported self medication with subtherapeutic doses of the analgesics, aspirin and paracetamol even though full doses of diclofenac was prescribed by the physician[45.]

The prevalence of self medication during pregnancy was low compared to the general population. This could be because drug use during pregnancy was mainly decided by the Obstetrician as revealed by the 5% prevalence rate from a multicentre study in Spain[46]. The rate was 10% among pregnant women of varying gestational ages in another Spanish study[47.]

Self medication and traditional medicine dominate alternative health care strategies of child health in the tropics[48]. However, this practice is not limited to the tropics as a study in Spain showed that of children faced with acute illness, 86.6% previously self medicated for respiratory symptoms[49]. Babies are not spared as large number are given "gripe water" for no valid reason or for only trivial symptoms by their mothers[50]. Self medication could also account for why some fail to complete their hospital treatment especially for chronic illness. It was reported that 72.9% of the non-attenders at paediatric tuberculosis out-patient clinic self medicated with the antituberculosis drugs intermittently and beyond the period allowable[51].

In patients with sexually transmitted diseases (STD), the prevalence of self medication might actually be higher than reported. A study in a STD clinic in the United States showed that while only 14% admitted self medication with antimicrobial agents, urinary assay was positive for 60% of those using the agents[52]. Failure to tell the truth on the questionnaire might be due to the stigma attached to their ailment. Urinary assay for household drugs was also used to determine drugs available for self medication from a survey of urban and rural households in Zimbabwe[53].

Malaria is one of the major killers in developing countries. The use of antimalarias was not free from self medication as revealed by a hospital-based study in Tanzania where 72.7% of patients reported having used home kept antimalaria medication for suspected malaria fever[54]. People can also self-treat for malaria using herbal remedies or medications purchased from local shops as a study shows that 60% of malaria cases were self-treated through this means while only 18% received treatment at the local health centre[55].

People afflicted with chronic illness sparingly see a Doctor for their ailments as they learn to cope using self medication. It was reported that nearly six million Americans with self-treated arthritis never saw a Doctor for their condition even with severe limitation of activit[56]. This was also seen among migraine sufferers in Kenya where a study revealed that 56% resorted to self medication though 40% sought medical attention[57].

Among commercial sex workers, self medication with antibiotics was perceived as a potential means of protection against STD and acquired immune deficiency syndrome – AIDS[58]. In smokers, the practice may be used to self-treat negative effects with nicotine as evidenced by the occurrence of major depression in some who try to quit the smoking habit[59].

Health care providers are favourably disposed to self medication. It was reported that General Practitioners expected other Doctors to self-treat themselves rather than consult their fellow colleagues[60]. Tong and co workers reported a 60% "ever used" rate of self medication among pharmaceutical representatives probably due to their continuous exposure to drug samples of pharmaceutical companies[61].

Self medication with antibiotics is a common practice. Of medications consumed for self treatment in Nigeria, it accounted for 63.4% in an urban slum[62], 44% among urinary tract attendees prior to hospital admission[63], 39% among medical undergraduates[64] and 24% for treating menstrual symptoms[65]. In addition, it accounted for a substantial percentage of diarrhoea treatment. A Nigerian study revealed that 53% of cases were self-treated with antibiotics while only 40% of cases were treated by prescriptions from the clinics[66]. The self-treated cases were usually associated with a higher risk of using inadequate medication or dosage. Bojali et al reported self medication with antibiotic for diarrhea in 37% of cases even though it is indicated in 5% of cases. It is noteworthy also that about 27% of cases used inadequate antibiotics in terms of duration and dose[67] though previous study reported 67.7%[68]. Among market women, self medication with antibiotics accounted for 18% of all drugs used for this practice[69]. However, 90.4% of cases had incorrect knowledge about its dose and duration[20]. Possible explanation for this high prevalence of incorrect dose had to do with the time constraint in following the six hourly regimes of antibiotics for at least five days. This may seem laborious once the symptoms abate compared to single daily drug dosages which antihelmintics, laxatives/purgatives and sedative/hypnotics are known for. This might account for the latter's correct dose which are easy to remember[20.]

Self medication with analgesics is a common practice. The prevalence rate among market women was 31.3% of all drugs used in self medication[20.] A population-based study in Sweden revealed that 35% used a form of analgesics in the past two weeks due to self-perceived poor health and pain[70]. A study among the disabled with painful ailments reveals that about 50% self medicated with analgesics everyday[71]. Majority of people with acute episodic headache self medicated with OTC analgesics which was believed to be more adequate than if prescribed while those with chronic headache treat themselves with prescribed drugs from previous doctor's visit[72].

In the dental profession, pain is the most likely symptom which could result in analgesic use without the Doctor's prescription. Dentists are aware that patients with dental pain often use OTC analgesics on their own to alleviate symptoms or to avoid the need for dental attendance altogether[29,33,73]. A study revealed that the current use-rate was 52.9% among children with post-operative dental pain[74]. Apart from pharmaceutical products used by the majority, a minority patients use dangerous substances to alleviate dental pain such as battery water, local gin and 'touch and go' solution[33], petrol and vinegar[29].

Drugs used for self medication in some countries are prohibited or strictly regulated in other countries. A study of some Mexican Pharmacies revealed that while 14.3% of drugs sold are strictly regulated, 51.4% of such drugs were obtained for self medication purposes[75]. Self medication with re-used needles and syringes for home injection of medications and vitamins may be a risk factor for transmission of HIV infection according to an exploratory study in the United States[76]. Apart from using prescribed drugs, natural medicines have also been used. For instance, 35% of women referred to a Gynaecologist admitted self medication with natural medicines[77]. Nutritional or dietary supplements like vitamins, minerals, herbal products, tissue extracts and protein solutions are also used by Americans as dietary supplements, for energy and immune system enhancement and cancer prevention[78].

Self medication has some life saving advantages. It has been shown that people self treating reflux oesophagitis with antacids had a low prevalence of pre-neoplastic and neoplastic pathologies while the use of alginate in 68% of cases relieves symptoms[79].

4. Self medication and drug dependency

The abuse of various self medication compounds for chronic illnesses may or frequently lead to a state of dependency. Aspirin, acetaminophen and caffeine were the most frequently abused among chronic headache sufferers[80]. Substance abuse and drug dependency have multiple causes ranging from poor instructions from the physician, improper diagnosis with gradual increase in amount consumed, a reinforcement mechanism and brain stimulation effects[80]. For instance, cocaine acts directly on the "pleasure centres" of the brain to release dopamine which triggers an intense craving for more of the drug otherwise a painful withdrawal symptom persist. It therefore produces pleasurable sensation of "reward" and physical dependence[81].

Nicotine, the psychoactive ingredient in cigarettes is an addictive agent that can stimulate and relax the user. Hence, some smokers self-treat negative moods with it[59]. Approximately, 30% of women from a study conducted in the United States, smoke cigarette during pregnancy despite its deleterious effect on the mother and foetus[82].The beverage, alcohol (ethanol) was so commonly consumed that it is seldomly thought of as a drug. When consumed in small quantity, it induces a feeling of well being and relaxation while in large amounts, intoxication is produced. It can therefore be used as a form of self medication to achieve any of these states[82]. It may also be used to cope with perceived problem of sexuality[83].

The relationship between self medication and drug dependency was explained with the self medication hypothesis of addictive disorders defined by Khantzian as motivation of patients to seek a specific drug (reinforcement mechanism) for relief of a particular set of symptoms for adaptive purposes[84]. However, not all cases of drug dependencies follow this hypothesis because there are traits or symptoms which separate various groups of drug dependent individuals[85,86]. As a result, Khantzian[87] revisited his theory in 2003 and stated that there was growing clinical support for the significant relationship between substance abuse disorders and psychiatric disorders as opposed to simple personality. Hence, people who are not receiving proper mental health treatment are attempting to selfmedicate for their disorders by using illicit substances.

5. Self medication – Sociodemographic and medical factors

Despite a growing research interest in self medication, little information has been available about its major determinants. Individual self care in illness is shaped in the social environment – a major determinant of the type and amount of health care services used[88]. The sociodemographic determinants are age, gender, occupation, education, marital status, religion, race, income and culture. The sociomedical factors may be related to the female reproductive role (pregnancy, breast feeding, and menstruation), psychiatric disturbance, medical states like asthma, migraine and so on.

The younger age group engaged in self medication than the older ones[29,44,56,70,89,90]. However, some studies revealed no association between age and self medication[16,24,91,92]. Women have above average knowledge about drugs and risks of self medication compared to men[19]. They also had a much higher probability of using supplements, OTC tranquilizers and analgesics for self medication than men who on the other hand commonly use more stimulants[3,70,93,94]. Self medication with drugs to relieve depressive symptoms was far more likely in men than women[95]. Factors related to general health status and women's reproductive role influences gender differences in self medication[96]. During breastfeeding, self medication was dictated by the mother and her infant's disorder. In addition, women with pre-menstrual symptoms use caffeine as a form of self medication to relieve the symptoms[97]. However, some studies revealed no association between gender and self medication[16,91,92].

Various studies consistently showed that self medication was associated with educational level. For instance, there is a positive correlation between level of education and self medication[16,18,24,65,89,98,99]. The trend of consulting patent medicine dealers for prescription decreases with acquisition of more formal education[24,98,100]. While studies showed no correlation between self medication and occupational status[17,18], others revealed some association. For instance, employment status affected the pattern of OTC and prescription drugs[96]. Specialist in anaesthesiology, emergency medicine, general and family practice self medication than other medical specialist probably due to habitual overwork and unrestricted access to drugs[101].

The relationship between race and self medication had been documented from various studies. Non whites had a higher probability of using tranquilizers than whites[94] and whites likely than blacks to consume supplements[93]. Among the elderly, fewer blacks reported the use of OTC medications than non-blacks[102]. While some studies found little or no association between self medication and social status[17], others reported that among school aged subjects, social classes of parents has a direct relationship with drug consumption among their children[103]. The influence of culture is common in health related states and was related to female reproductive roles like childbirth, and in the treatment of morbidity and mortality in children[104]. Athletes consume sex hormones to alter their menstrual cycle so as not to disturb the training schedule and competitive programme while some use anabolic steroids to enhance their performance[10].

6. Self medication – Commonest complaint responsible

Usually, self medication is indicated for trivial symptoms perceived by the patient. It was favoured for skin condition, general health care, aches and pain, problems of the eye, mouth, gastrointestinal and respiratory tract[105]. Among adult patients with acute pathology, the most common complaints were pain and increase body temperature[106]. In a recent rural population study in Nigeria, it was in the order: malaria, gastrointestinal problems and urinary tract infections[107]. Among rural Japanese housewives it was headache, tiredness and gastrointestinal problems while in American and British housewives, it was emotional or psychological complaints[108]. Among children, respiratory symptoms [49] especially for common cold [92] with or without fever [109] were the commonest complaints.

With the use of antibiotics, the indication varies with different studies. The commonest complaints were for soft tissue, sexually transmitted diseases, upper respiratory and gastrointestinal tract infection[110;] upper respiratory tract infection[91;] respiratory infection[38]; throat, dental and urogenital infection[111]; respiratory tract infection and acute diarrhoea[68,112] and diarrhoea, abdominal pain, fever and rashes[20].

For analgesics, the commonest complaints associated with its use are limb and back pain[45;] self perceived pain and poor health[70]; and body pain, headache, body weakness and fever[5]. Supplements are consumed for enhancement of diet, energy, immune system and for cancer prevention[78].

7. Self medication – Commonly used medications

Several medications have reportedly been used for this practice. This included antibiotics, analgesics and vitamins[1,] analgesics, vitamins and oral antibiotics among primary care patients[16], while for OTC drugs, the commonly requested were for nervous system[113], analgesics, cough or cold medications[42]. Among adult married women, the commonly used medications were vitamins and contraceptives[114].

Among market women surveyed in a sub-urban community in Nigeria, antipyretic analgesics, haematinics/vitamins, antibiotics, antimalarials and alternative or traditional medicines respectively were commonly consumed[69]. In an European study of those presenting with acute illness, the most commonly used medications were analgesics and antipyretics[106] and among paediatric presentations were antipyretics, analgesics, antitussives and antibiotics[49]. In a community-based pharmacy study in Portugal, the main therapeutic groups used for self medications were in the order: throat, cough, cold, stomatological, laxative, analgesics and dermatological products respectively[54]; antibiotics and antimalarials for illness management[107] and analgesics and antibiotics in dental outpatients[20] from recent Nigerian studies; analgesics, cough, cold remedies, antiallergies, vitamin and energy tonic were the commonest OTC used as revealed from a recent review of selfmedication in India[89].

Orthodox medications were preferred to traditional African medicines for most common symptoms. However, some studies in developing countries revealed that people prefer traditional African medicines for diarrhea, vomiting, cough and cold[115], rheumatic and neurological complaint[100]. Among Hong Kong Chinese, Chinese tonic was the most frequently used traditional medicine for self medication which was perceived as equally effective as western medicine[21]. The most commonly used supplement among Americans were minerals, multivitamins, vitamin C, calcium, vitamin E and A[93] while the remaining percentage were for herbal products, megadose vitamins, protein and amino acid preparations[78.]

8. Self medication – Places where drugs are obtained and sources of drug knowledge

The common places for drug supply were in the order: pharmacies, general medicine dealers, hospital/clinics, traditional sources, private practitioners and other sources[115] like household medicine cabinet containing previous medical prescriptions which may not have been prescribed for the same condition[116]. Recent studies agreed that the pharmacy, roadside/patent medicine stores were the commonest places where drugs were obtained for selfmedication purposes.

Other studies[90,117] also agrees that family medicine cabinet were sources of self medication. However, the common sources of household stock are chemist, pharmacy, supermarket, hospital/clinics, friends and relatives[118]. In developing countries common sources of antimalarials used for self treatment were street and village shops and this could account for up to half of antimalarial drug distribution[119.] Recent studies in Nigeria and Cameroun[24,29] agrees with the above showing that the hospital/pharmacy, patent /road side medicine dealers, and local hawkers/mobile drug vendors and native healers were the commonest sources.

In choosing the most appropriate medicine to buy from the chemist shop, people relied on the advice of the sales clerk in the chemist shop[1] , print media, family and friends[24,29,120,121], pharmacist, general medicine dealers, general and private medical practitioners[24]. Among the young ones, sources of drugs knowledge include family members especially the mother (for therapeutic purposes), peer groups and illegal market (for intoxication purposes) [122]. Among secondary school pupils in an Hong Kong study, the sources were in the following order: family members, previous illness experience, pharmacy shops, doctors or nurses, television or radio, newspaper or magazines, friends and teachers[90].

For painful condition, people self select drugs for self medication while small percentages were advised by the pharmacist or non-health professionals[65,123]. Since individuals suffering from sexually transmitted diseases often treat themselves with antibiotics, the common sources of drug supply were the medicine cabinet at home and the sources of drug knowledge were family members and friends. For dietary supplements, the principal source of drug information was the mass media[52].

9. Self medication – How and when

People who self medicated reported taking one or several medications and more often one or two medications were involved[20,26,33,116]. Individuals sometimes self administer medications via drug identification. Trade names were common means of identification and less frequently by generic names, action, color, shape and common usage names[24,124].

In painful complaints, the number of analgesics and duration of consumption was directly related to the intensity of the pain[123]. This was collaborated by a study among dental patients which revealed that the majority use analgesics within one week of presentation and only present when the pain did not resolve[20]. Self medication is commonly associated with subcurative doses. This can be seen from antimalaria therapy with chloroquine either administered orally or via injection[98] and with antibiotic use where two-third of individuals used it for less than five days or in insufficient quantity[20,68,112].

Among pregnant women interviewed, over fifty types of symptoms necessitated self treatment[47]. In asthmatics, most of the sufferers tended to reduce their doses of medication following improvement of their symptoms[39]. Sometimes, consumption of household medications may be incorrectly volunteered but could only be confirmed by urine screening test for such medications[53].

10. Self medication – Side effects and risk

Although these medications are considered risk free and useful for the treatment of common health problems, their excessive use can also lead to serious side effects and unfavourable reactions[89]. For instance, the therapy may be poorly suited for the illness in question, delay diagnosis and the beginning of effective therapy, increased inorganic risk(s) due to inadequate drug therapy or of unnecessary expense[116] and drug interaction between prescription and non prescription drugs[99].

The prevalence of side effects was associated with lack of knowledge about the drug prior to its usage[21]. Insufficient curative treatment with chloroquine (CQ) for individuals who treat themselves for suspected malaria fever could result in resistance to Plasmodium falciparum – the agent causing the ailment[125,126]. Chronic CQ toxicity was important in the causation of heart block in Africa, CQ retinopathy and abnormal ophtalmological findings, cardiac arrhythmia[127,128]. Stevens–Johnson syndrome following self medication with Fansidar has been reported[129.]

With respect to OTC medications, reported risks associated with the improper use includes addiction, gastric irritation, liver toxicity, rebound headache syndrome[41] , milk alkali syndrome[130]; dental caries from prolonged usage of self administered mineral supplement containing lactose[131]; liver toxicity/failure following prolonged use of analgesic containing paracetamol for dental pain[132] ; peripheral neuropathy and subdermal vascular dermatosis following Vitamin B6 megatherapy[133]; cholinergic excess, loss of consciousness and seizure following cutaneous application of Diazinon, an organophosphate insecticide for pubic lice[134]. In addition, laxative abuse causing ammonium renal urate calculi, gastrointestinal fluid and electrolyte loss resulting in chronic extracellular volume depletion and intracellular acidosis had also been reported[135]. Simbi et al recently reported in-utero-ductal closure following near term maternal self medication with Nimesulide and Acetaminophen[136.] Self administered oral diuretics could result in pseudo-barter syndrome (hypokalaemia, metabolic alkalosis, hyperaldosteronism, hypomagnesimia, normocalcimia and hypocalcuria)[137]. Topical anaesthetic abuse of the cornea with subsequent fungal (candida) keratitis[138] and severe toxic keratopathy[139] had been reported.

Sometimes, the side effect which could be dermatological tends to be the primary cause of drug intolerance. For instance, cutaneous manifestation of psoriatic arthritis could be exacerbated with ibuprofen self therapy[140], fixed pigmented eruptions could be manifestation of such drugs, which if unrecognized, might be fatal if such a drug was repeated[141].

Among the elderly, adverse reaction to drugs are characteristically more frequent and severe as a result of factors including self medication[142]. In the case of substance abuse, depending on the substance used, it may result in organ damage, medical complications, vascular injury, less than satisfactory quality of life and depression[80]. Among alcoholics, male and female fertility can be interfered with[82].

Drug use before hospital admission is a source of potential drug toxicity and may obscure the diagnosis of infective illness and delay hospital stay[143]. It has been shown that the five most common adverse events following self medication related hospitalization were upper gastrointestinal bleeding[144], skin rashes, hypoglycaemia, hypercorticism and hepatitis[145].

11. Reasons for self medication

The common reasons could be to cure an ailment[24], suppress its cause indefinitely to give the body time to completely overcome it or for prevention, prophylaxis, palliation, convenience, postponing a natural event, out of habit or for special purposes[10.] In some cases, the main reasons could be triviality of the symptoms[24,25,105], to save money and time[16,24,33,120], lack of gravity to go and see a physician because they can take care of themselves[117] or previous medical prescription for related symptoms[21,109.]

In specific diseases like acute non specific diarrhea, people self medicated because the ailments were of short duration, can be treated symptomatically with non-prescription medications and adequate hydration and do not require a visit to the physician office[146]. In the case of chronic illnesses, it could be the cost of medication, patient's psychological status, perceptions of the seriousness of their illness and vulnerability to complications[147]. For antimalarials, self medication with orthodox medication was greater than traditional remedies because of their efficiency, popularity, cheapness and availability[98], distance and cost of seeking care from the formal health service[99] and cultural beliefs[148]. Among market women, reasons given for self medication was in the order: for minor ailments, cheapness and because they know what to do[24].

Among dental patients, since the commonest complaint was pain, the main reason for self medication was to serve as a means of avoiding the need for dental attendance altogether[73]. Acute headache sufferers may treat themselves with OTC if they perceive it to be more adequate than prescribed drugs[72]. Reasons for using psychoactive drugs among the young people range from insomnia, worry or depression to intoxication[122] while smokers may self treat negative effects like major depression with nicotine[59].

12. Self managed health care

Self medication is a necessary and important aspect of daily health care. It encourages self reliance for curative, preventive, promotive and rehabilitative care[18]. It appears to be substitute for, rather than supplements or stimuli for health service utilization[149]. In the Federal Republic of Germany and Switzerland, its importance in health care system had been recognized because of possibility of self treatment of minor illness and its health economic benefits[150,151].

Since individuals have a certain right to reasonable self mediation, an important aspect of a qualitative improvement of the practice was the information, education and counseling of the patient of which the pharmacist plays a major role[152]. In view of this, Ruegg reported that pharmacist in Switzerland had accepted this aspect of patient's education and are adjusting their education to the problems of self medication[150]. This role of Pharmacist had also been suggested in a later study[24,153,].

In some cases, the practice is frequently and successfully used. An Australian-based study revealed that in only 2% of cases self treatment for minor ailments were the actions taken assessed as inappropriate and potentially harmful[154]. This agrees with a later study which showed that few, if any were consuming nutrient supplements in amount considered toxic[93] and that most consumers used self medication preparations in a safe and proper way[155]. This agrees with other studies[20,109]. Hence, in some patients, self medication was recommended if they continue to have recurrences of a chronic infective process[156]. Further, because OTC drug sales statistics reflects pattern of self medication, it may be used to monitor the practice[4].

The above reflects the need for a liberal regulatory environment and comprehensive information package in consumer-oriented language. This could be achieved via consumer-oriented advertisement and consumer product package leaflet. Advertisement gives consumers choice to determine what to buy. The government benefits since the consumers can buy OTC drugs with their own money and does not engage government health care budget for minor ailments[157]. Hence, one potential means of reducing pressure on the health budget of a country's gross domestic product (GDP) is a greater reliance on self health care[7]. In a rehabilitative setting, patients could be actively involved in their medication program and be independent on the use of their medications when they leave hospital. A self medication program fulfills this role[158,159].

Therefore, during drug advertisement, advertising agencies should emphasize the possible side effects as they do for cigarette smoking. By this people are well informed as they read or hear it (especially if illiterate). Because the practice of self medication is worldwide, careful drug history by General Practitioners and Physicians is important to know what patients are taking before treatment commences especially when subtherapeutic doses are involved.

13. Conclusion

Self medication is a necessary and important aspect of primary health care which if properly managed could be incorporated in the health care delivery system to reduce the burden on the secondary and tertiary level so that attention could be focused on the more serious health problems.

14. References

[1] Haak H. Pharmaceuticals in two Brazillian Villages: Lay practices and perceptions. *Soc. Sci. Med.* 1988; 27(12): 1415-1427

[2] Menard G., Allain H., Le Roho S., Morel G., Beneton C. A.single day survey in pharmacies on the consumption of analgesics and antipyretics. *Therapie* 1993; 48(3): 263-267.

[3] Casner P.R and Guerra L.G. Purchasing prescription medication in Mexico without prescription: The experience at the border. *West J Med* 1992; 156(5): 512-516

[4] Wessling A. Over-the-counter sales of drug in Sweden 1976-1983. *Eur J Clin Pharmacol* 1989; 33(1):1-6.

[5] Russel JM, Baton SE, Lawrence AG. Self medication by women attending a genitourinary medicine clinic. *Int J STD AIDS*. 1990; 1:279-281.

[6] Gordon SM, Mosure DJ, Lewis J, et al. Prevalence of self medication with antibiotics among patients attending a clinic for treatment of sexually transmitted diseases. *Clin Infect Dis.1993*; 17:462-465.

[7] Blenkinsopp A. Cost benefit of self prescribing. *Lancet* 1989; 1(8651):1393.

[8] Johnson L.W. The demand for OTC medicines: Some Australian evidence. *Asia Pac. J Public Health* 1991; 5(3): 228-235.

[9] Mississippi Weekly Report. Self treatment with herbal and other plant derived remedies in rural Mississippi 1993. MMWR *Morb. Mortal. Wkly. Rep.* 1995; 44(11): 204 -207.

[10] Olatunde A. Self medication: Benefits, Precautions and Dangers. 1st ed., Macmillan Press Ltd, 1979; 6-13.

[11] Helman C. Culture, Health and Illness. An introduction for Health Professionals.2nd Ed., Butterworth-Heineman Ltd, 1991; 54-56

[12] Abrahams N., Jewkes R., Mvo Z. Indigenous healing practices and self medication among pregnant women in Cape Town, South Africa *Afr. J. Repr. Health* 2002; 6(2): 79-86.

[13] Cutting W. Self prescribing and promotion of antidiarrhoea drugs. *Lancet* 1989; 1 (8646): 1080-1081.

[14] Tupasi T.E., Miguel C.A., TalloV.L., Bagasao T.M., Natividad N., et al. Child care practices of mothers: implications for intervention in acute respiratory infections. *Ann Trop. Paediatr.* 1989; 9(2): 82-88

[15] Wegner P.J., Philips W., Radford M., Hornsby J.L. Frequent use of medical services. Patient reports of intentions to seek care. *Arch. Fam. Med.* 1995; 4(7): 594-599.

[16] Saeed A. Self medication among primary care patients in Faradak clinic in Riyadh. *Soc. Sci. Med.* 1988; 27(3): 287-289.

[17] Szyllejko O. The use of in prescribed medicine in Warsaw during the years 1970-1980. *Drug Intell Clin Pharm.*1984; 18(9): 494-507.

[18] Abosede O. Self medication: an important aspect of primary health care. *Soc.Sci. Med.* 1984; 19(7): 699-703.

[19] May U. Status of drug knowledge in the population. *Soz. Praventimed* 1986; 31 (3):151-155

[20] Afolabi AO. Drug Use and Misuse in an Adult Nigerian Population. *Clinical Review Journal* 2007; 68: 7-14.

[21] Lam C.L., Catarivas M.G., Munco C., Launder I.J. Self medication among Hong Kong Chinese. *Soc. Sci Med.* 1994; 39(12): 164-167.

[22] Awad A, Eltaved I, Matowe L, et al. Selfmedication with antibiotics and antimalarials in the community of Khartoum State, Sudan. *J Pharm Sci.* 2005; 8(2): 326-31.

[23] Bamgboye EA, Amoran OE, Yusuf OB. Self medication practices among workers in a tertiary hospital in Nigeria. *Afr J Med Med Sci* 2006; 35(4): 411-5.

[24] Afolabi AO. Factors Influencing the Pattern of Self-Medication in an Adult Nigerian Population. *Ann Afr Med* 2008; 7(3): 120-127.

[25] Omolase CO, Afolabi AO, Mahmoud A.O, Omolase B.O. Ocular self-medication in Owo, Nigeria. *Nig. J Postgrad Med* 2008; 1(1): 8-14.

[26] Omolase C.O, Adeleke EO, Afolabi AO, Afolabi OT. Self medication amongst general outpatients in a Nigerian community. *Annals of Ibadan Postgraduate Medicine* 2007; 5(2): 65-68.

[27] Servidoni A.B, Coelho L, de Lima Navarro M, et al. Self-medication profile of ENT patients. *Rev Bras Otorrinolaringol* 2006; 72(1): 83-8.

[28] Onajole A.T, Bamgbala A.O. Socio demographic characteristics of drug misuse in a Polytechnic in Lagos, Nigeria. *Nig Jnl Health and Biomed Sciences* 2004; 3(1): 40 -43.

[29] Agbor MA, Azodo CC. Self medication for oral health problems in Cameroon. *Int Dent J 2011*; 61(4); 204–209.

[30] Awad A.I, Eltayeb I.B, Capps P.A. Selfmedication practices in Khartoum State, Sudan. *European Journal of Clinical Pharmacology* 2006: 62(4): 317 – 324.

[31] Awad A.I, Eltayeb I.B. Self-Medication Practices with Antibiotics and Antimalarials among Sudanese Undergraduate University Students. *Ann Pharmacother* 2007; 41(7): 1249-55.

[32] Hartlova S, Solich J. Drugs and health awareness in the population. *Cesk Zdrav* 1990; 38(3): 120-126.

[33] Afolabi AO, Akinmoladun VI, Adebose IJ, Elekwachi G. Self-Medication Profile of Dental Patients in Ondo State, Nigeria. *Nig J Med* 2010; 1: 96-103.

[34] Abula T and Worku A. Self medication in three towns of North West Ethiopia. *The Ethiopian Journal of Health Development* 2001; 15 (1) 25 – 30.

[35] Ajaiyeoba A.I and Scott S.C.O. Risk factors associated with eye diseases in Ibadan, Nigeria. *African Journal of Biomedical Research* 2002: 5(1-2): 1 – 3.

[36] Afolabi OA, Ehalaiye BF, Fadare JO, Abdur-Rahman AB, Ehalaye DN. Survey of ototopical self medication among patients attending ENT and family medicine departments in a Nigerian hospital. *Eur J Gen Prac* 2011;17(3); 167-170.

[37] Osinusi K and Oyejide C.O. Child care practices with respect to acute respiratory tract in a poor urban community in Nigeria. *Rev. infect. Dis.* 1990; 12(supp 18): 1039-1041.

[38] Verbrugge L.M and Ascione F.J. Exploring the iceberg. Common symptoms and how people care for them. *Med. Care.* 1987; 25(6): 539-569.

[39] Pretet S., Perdrizet S., Poisson N., Pujet J.C., Marsac J. Treatment compliance and self medication in asthma in France. *Eur Respir J.* 1989; 2(4): 303 -307

[40] MacGregor E.A., Vohrah C., Wilkinson M. Analgesic Study: a study of treatments used by patients for migraine prior to attending the city of London Migraine Clinic. *Headache* 1990; 30(9): 571 – 574.

[41] Robinson R. Pain relief for headaches. Is self medication a problem? *Can Fam Physician.* 1993; 39: 867-868, 871-872.

[42] Price L.J. In the shadow of biomedicine: Self medication in two Ecuadorian Pharmacies. *Soc. Sci. Med.* 1989; 28(9): 905 – 915.

[43] Lansang M.A., Lucas-Aquino R., Tupasi T.E., Mina V.S., Salazer L.S., et al. Purchase of antibiotics without prescription in Manila, the Philippines. Inappropriate choices and doses. *J Clin Epidemiol* 1990; 43(1):61-67.

[44] Martins A.P., Miranda-Ada C., Mendes Z., Soares M.A., Ferreira P., et al. Self medication in a Portuguese urban population: a prevalence study. *Pharmacoepidemiol Drug Sef* 2002; 11(5): 409-414.

[45] Bosch F., Toranzo I., Banos J.E. A survey of pain complaints and treatment by General Practitioners in a Spanish public health organization. *Clin J Pharm* 1990; 6(3): 206 – 211.

[46] DUP Workshop in Spain: Multicenter study of the use of drugs during pregnancy in Spain (111). Drugs used during the first pregnancy trimester. *Med Clin Barc* 1991; 96(2): 52 – 57.

[47] Ortigosa-Corona E., Carrasco-Reseudiz I., Gonzalez-Flores A., Damaso-Ortiz M. Gestational Disorders. *Gyneol Ostet Mex* 1993; 61: 247-253.

[48] Assimadi K., Assogba L., Locoh-Donou T., Ategbo S., Siamevi K., et al. Routes of care and health of children in Togo. *Bull Soc Pathol Exot Filiales* 1991 ; 84 : 794-803.

[49] Oriol-Toron P.A., Lou Arnal S., Blasco-Perez-Aramendia M.J., Sadiles Cabello A.I., Perez-Ramirez I. Self healthcare when faced with acute patholgy in childhood. *Aten Primaria* 1994; 14(2): 616 – 618.

[50] Illingworth C and Timmins G. Gripe water: what is it? Why is it given. *Health Visit* (1990); 63(1): 378.

[51] Oviawe O and Ojemudia E. The problem of non-attendance at a paediatric tuberculosis outpatient clinic. *Ann Trop Paediatr* 1993; 13(3): 243 – 247

[52] Gordon S.M., Mosure D.J., Lewis J., Bron S., McNagny S.E., et al. Prevalence of self medication with antibiotics among patients attending a clinic for treatment of sexually transmitted diseases. *Clin Infec Dis* 1993; 17(3): 462-5.

[53] Stein C.M., Gora N.P., Macheka B.M. Self medication in urban and rural Zimbabwe communities. *Eur J Clin Pharmacol* 1989; 27(6): 741-747.

[54] Mnyika K.S., Killewo J.Z., Kabalimu T.K. Self medication with antimalaria drugs in Dares-Salaam, Tanzania. *Trop Georg Med* 1995; 47(1): 32-34.

[55] Ruebush T.K., Kern M.K., Campbell C.C., Oloo A.J. Self treatment of malaria in a rural area of Western Kenya. *Bull World Health Organ* 1995; 73(2): 229-236.

[56] RaO J.K., Callahan L.F., Helmic C.G. Characteristics of persons with self reported arthritis and other rheumatic conditions who do not see a Doctor. *Am J Rhum* 1997; 24(1): 169-173.

[57] Amayo E.O., Jowi J.O., Njeru E.K. Migraine headaches in a group of medical students at the Kenyatta National Hospital, Nairobi. *E Afr Med Jour* 1996; 73(9): 594-597.

[58] Abellonosa I and Nichter M. Antibiotic prophylaxis among commercial sex workers in Cebu city, Phillipines. Pattern of use and perception of efficacy. *Sex Trans Disease* 1986; 23(5): 407-412.

[59] Bock B.C., Goldstein M.G., Marcus B.H. Depression following smoking cessation in women. *J Subs Abuse* 1996; 8(1): 137-144.

[60] Chambers R. What should Doctors do if they become sick? *Fam Pract* 1993; 10(4): 416-423.

[61] Tong K.L and Lien C.Y. Do Pharmaceutical representatives misuse their drug samples? *Can Fam Physician*. 1995; 41: 1363-1366.

[62] Kehinde OO, Ogunnowo BE. Pattern of Antibiotic use by the Public in an urban slum.http://journalmanagersorg/files/journals/1/articles/22/submission/original/22-67-1-SM.doc.

[63] Enwere OO, Agina US. Self medication as a factor for antibiotic resistance of urinary pathogens in hospitalized medical patients. *Nigerian Journal of Medicine* 2011 [epud ahead of print]

[64] Fadare JO, Tamuno I. Antibiotic self-medication among university medical undergraduates in Northern Nigeria. *J Public Healt Epidemiol* 2011; 3(5): 217-220.

[65] Sapkota AR, Goldstein RE, Coker ME, Sweet SJ, Sopeju PO, Ojo MT et al. Self-medication with antibiotics for the treatment of menstrual symptoms in southwest Nigeria: a cross-sectional study. *BMC Public Health* 2010; 10:610doi:10.1186/1471-2458-10-610.

[66] Oni G.A., Schumann D.A., Oke E.A. Diarrhoeal disease. Morbidity, risk factors and treatment in a low socio-economic area of Ilorin, Kwara State, Nigeria. *J Diarrhoel Dis Res.* 1991; 9(3): 250-257.

[67] Bojalil R. and Calva JJ. Antibiotic misuse in diarrhoea. A household survey in a Mexican community. *J Clin Epidemiol* 1994; 47(2): 147-156.

[68] Bojalil R., Calva J.J., Ortega H., Aguaron-Joven E., Adan-Gil F.M., Pons-Pons L., et al. The use of antibiotics in a community of Mexico City. A household survey. *Bol Med Infant.* 1993; 50(2): 79-87.

[69] Afolabi AO, Ojo MA. Common medications consumed by Market Women in a suburban Community In Lagos, Nigeria. *Tropical Journal of Health Sciences* 2009; 16(2): 1-5.

[70] Antonov K and Isaacson D. Use of analgesics in Sweden. The importance of sociodemographic factors, physical fitness, health and health related factors, and working condition. *Soc. Sci. Med.* 1996; 42(11): 1473 – 1481.

[71] Astin M., Lawton D., Hirst M. The prevalence of pain in a disabled population. *Soc. Sci. Med.* 1996; 42(11): 1457-1464.

[72] Ferarri A., Stefani M., Sternieri S., Bertolotti M., Sternieri E. Analgesic drug taking: belief and behaviour among headache patients. *Headache* 1997; 37(2): 88-94.

[73] Preshaw P.M., Meechan J.G., Dodd M.D. Self medication for the control of dental pain: What are our patients taking. *Dent Update* 1994; 21(7): 299-301, 304.

[74] Asc G and Drazner E. The incidence of post-operative and analgesic usage in children. *ASDC J Dent Child* 1992; 59(1): 481-52.

[75] Vincencio-Acevado D., Alfaro-Valle A., Martinez-Toledo J.L. Characteristics of drug acquisition in Morelia (Michoaclian), Mexico. *Bol Officina Sanit Param* 1995; 119(3): 236-242.

[76] Flaskerud J.H and Nyamathi A.M. Home medication injection among Latina women in Los Angeles. *AIDS Care* 1996; 8(1): 95-102.

[77] Kristoffersen K and Thomsen B.W. Use of natural medicines in women referred to specialists. *Ugeskr for Laeger* 1997; 159(3): 294-296.

[78] Eliason B.C., Myszkowski J., Marbella A., Rasmann D.N. Use of dietary supplements by patients in a family practice clinic. *J Am Board Fam Prac* 1996; 9(4): 249-253.

[79] Corder A.P., Jones R.H., Sadler G.H., Daniels P., Johnson C.D. Heartburn, Oesophagitis and Barret oesophagus in self medicating patients in general practice. *Br J Cli Prac* 1996; 57(5): 245-248.

[80] Elkind A. Drug abuse and headache. *Med Clin North Am* 1991; 75(3): 717-732.

[81] Office for substance abuse prevention. United States Department of Health and Human Services. What you can do about drug use in America. *DHHS Publication* 1991; No ADM 91-1572: 2,3,5-7.

[82] Cook P.C., Peterson R.C., Moore D.T. Alcohol, Tobacco and other drugs may harm the unborn. *DHHS Publication* 1993; ADM 90-1911: 2,4,19,20,23.

[83] Roman P. Biological features of woman's alcohol abuse. A review. *Public Heath Rep* 1988; 106(8): 628-637.

[84] Khantzian EJ. The self medication hypothesis of addictive disorders: Focus on heroin and cocaine dependence. American Journal of Psychiatry 1985; 1259-1264.

[85] Castaneda R., Lipshutz-Galanter M., Franco H. Emperical assessment of the self medication hypothesis among dually diagnosed in-patients. *Compr Psychiatry* 1994; 35(3): 180-184.

[86] Schinka J.A., Curtis G., Mulloy J.M. Personality variables and self medication in substance abuse. *J Pers Assess* 1994; 63(3): 413-422.

[87] Khantzian EJ. The self medication hypothesis revisited: the dually diagnosed patient. Primary Psychiatry 2003; 10: 47-48, 53-54.

[88] Deak K. Lay care in illness. *Soc Sci Med* 1986; 22(2): 125-130.

[89] Jain S. Concept of Self Medication: A Review. *International Journal of Pharmaceutical & Biological Archive* 2011; 2(3): 831-836.

[90] Tse M.H., Chung J.T., Mungo J.G. Self medication among secondary school pupils in Hong Kong. A descriptive study. *Fam. Pract.* 1989; 6(4): 303-306.

[91] Pucynski M.S., Gonzalez J., O'Keefe J.P. Self treatment with antibiotics. *Soc Sci Med* 1989; 28(9): 905-915.

[92] Rodriguez-Benito U., Magro-Peteguer R., Masip-Lopez M., Vacas Gamido R., Urbano-Rodriguez B. Self medication in general paediatrics. *Atn Primaria* 1994; 13(6): 313-316.

[93] Subar A.F and Block G. Use of Vitamin and mineral supplements: demographics and amounts of nutrients consumed. The 1987 Health Interview Survey. *Am J Epidemiol* 1990; 132(6): 1091-1101.

[94] Bell R. Over the counter drugs: factors in adult use of sedative, tranquilizers and stimulants. *Public Health Rep* 1984; 99(3): 9-23, 31.

[95] Weiss R.D., Griffin M.L., Mirin S.M. Drug abuse in self medication for depression. An emperical study. *Am J Drug Abuse* 1992; 18(2): 121-129.

[96] Zadoroznyj M and Svarstad B.L. Carme B., Dhelloh H., Senga J., Nzingoula S., Plassard H., et al. Clinical Presentation of non-pernicious malaria attacks in patients hospitalized in Brazzaville, Congo. *Bull Soc Pathol. Extol Filiales* 1991; 84(3): 266 – 272.

[97] Rossignol A.M., Bonnlander H., Song L., Phillips J.W. Do women with premenstrual symptoms self medicate with caffeine? *Epidemiology* 1991; 2(6): 403-408.

[98] Ezedinachi E.N., Ejezie G.C., Emeribe A.O. Problems of chloroquine resistant P.falciparum in Nigeria: One antimalaria drug utilization in metropolitan Calabar. *Cent Afr J Med* 1991; 37(1): 16-20.

[99] Habeeb G.E Jr and Gearhart J.G. Common patient symptoms. Pattern of self treatment and prevention. *J. Miss State Med Assoc* 1993; 34(6): 179-181.

[100] Carpetier L., Prazuck T., Vincent-Ballereau F., Ouedraogo L.T., La Faix C. Choice of traditional and modern treatment in West Burkina Faso. *World Health Forum* 1995; 16(2): 198-202.

[101] Chambers R. Health and Lifestyle of General Practitioners and Teachers. *Occup Med* 1992; 42(2): 69-78.

[102] Hanlon J.T., Fillenbaum G.G., Burchett B., Wall W.E Jr., Service C., et al. Drug use patterns among black and non black community dwelling elderly. *Ann Pharmacother* 1992; 26(5): 679-685.

[103] Lapeyre-Mestre M., Lebret-Bories E., Charlet J.P., Montastruc J.L., Montastuc P. Consumption of drugs in a population of school age subjects: *Therapie* 1991; 46(1): 49-53

[104] Basu A.M. Cultural influences on health care use: two regional groups in India. *Stud Fam Plann* 1990; 21(5): 275-286.

[105] Ibrahim M. Treating one's own ailment. *World Health Forum* 1996; 17(4): 409-410.

[106] Clenchet T., Navazo I., Llor C., Llovet D., Vilo A., et al. Self care measures recorded for acute pathology in primary care. *Aten Primaria* 1993; 11(4): 165-166, 168-169.

[107] Enato EFO, Sounyo AA,, Einarson TR. Medication Utilization and Illness Management Study in Nigeria. *The Annals of Pharmacotherapy* 2011; 45 (7-8): 924-930.

[108] Tonai S., Maezawa M., Kamei M., Satch T., FukuiT. Illness behaviour of housewives in a rural area in Japan: a health diary study. *Cult Med psychiatry*1989; 13(4): 405 -417

[109] Moral-Serrano S., Aguaron-Joven E., Adan-Gil F.M, Pons-Pons L., Baquer-Masgrau A., et al. Do the patients medicate themselves correctly? *Aten Primaria* 1994; 13(5): 242-246.

[110] Obaseki-Ebor E.E., Akerele J.O., Ebea P.O. A survey of antibiotic out-patient prescribing and antibiotic self medication. *J Antimicro Chemother* 1987; 20(5): 759-763.

[111] Ausejo M., Apechechea C., Rodriguez M.J., Salgado L., Puerta M.C., et al. A study of anti-infective agent. Self medication in two pharmacy offices. *Aten Primaria* 1993; 11(1): 41-43.

[112] Calva J. Antibiotic use in a peri-urban community in Mexico and drug store survey. *Soc Sci Med* 1996; 42(8): 1121-1128.

[113] Russel J.W., Barton S.E., Lawrence A.G. Self medication by women attending a genitourinary medicine clinic *Int. J STD-AIDS* 1990; 1(4): 279-281.

[114] Woods N. Self care practices among young adult married women. *Res Nurs Health* 1985; 8(3): 227-233.

[115] Joubert P.H., Sebata P.D., Van Reenen O.R. Self medication in a developing community. *S Afr Med J* 1984; 65(4): 129-131.

[116] Queneau P., Deco-Usus H., Jourlin M, Laurent H., Perpoint B. Self medication in rheumatologic consultation. Apropos of a study of 895 patients. *Rev Rheumal Osteoartic* 1985; 52(2): 79-84.

[117] Hugues F.C., Jeunne C., Saubadu S., Eme D., Denormandie P. A survey of self medication. Comparison of results obtained at two centres. *Therapie* 1990; 45(4): 325-329.

[118] Kasilo O.J., Nhachi C.F., Mutangadura E.F. Epidemiology of household medications in urban Gweru and Harare. *Cent Afr J Med* 1991; 37(6): 167-171.

[119] Foster S. Pricing, Distribution and use of antimalaria drugs. *Bull World Health Organ* 1991; 69(3): 349-363.

[120] Wen Y, Lieber E, Wan D, Hong Y. A qualitative study about self-medication in the community among market vendors in Fuzhou, China. *Health & Social Care in the Community* 2011; 19(5): 504-513.

[121] Ranno B.S. What characterizes elderly women who use over-the-counter vitamin and mineral supplements. *J Am Diet Assoc* 1988; 88(3): 347-348.

[122] Pederson W. Young people's use of psychopharmaceuticals. Self medication and intoxication. *Tidsskr Nor Laegeforen* 1989; 107(17-18): 1905-1908.

[123] Banos J.E., Bosch F., Toranzo I. Self medication with analgesics. A study on odontalgia. *Med Clin Barc* 1991; 96(7): 248-251.

[124] Yelland M.J and Vietch P.C. How do patients identify their drugs. *Aust Fam Physician* 1989; 18(11): 1441-1445.

[125] Sansonetti P.J and Charmot G. Chemical resistance of plasmodium falciparum in Africa. Current situation, implications for chemoprophylaxis. *Biomed Pharmacother* 1987; 41(4):167-171.

[126] Fowler V.J Jr., Lemnge M., Irare S.G., Melacela E., Mhina J., et al. Efficacy of Chloroquine Plasmodium Falciparum transmitted at Amani Eastern Usambara Mountains, North-Eastern Tanzania, an area where malaria has recently become endemic. *J Trop Med Hyg* 1993; 96(6): 337-345.

[127] Ihenacho H.N and Magulike E. Chloroquine abuse and heart block in Africans. *Aust NZJ Med* 1989; 19(1): 17-21.

[128] Demaziere J., Fourcade J.M., Busseuil C.T., Adeleine P and Meyer S.M., et al. The hazards of chloroquine self prescription in West Africa. *J Clin Toxicol* 1995; 33(4): 369-370.

[129] Phillip-Howard L. Prophylactic self-treatment with Fansidar resulting in Stevenson-Johnson syndrome in a five year old English boy. *Lancet* 1989; 2(8660): 803-804.

[130] Walsh H. The preventive management of extensive caries induced by self administered medications: A case report. *Clin Prev Dent* 1989; 116(6): 17-20.

[131] French J.K., Holdaway I.M., Williams LC. Milk alkali syndrome following over-the-counter antacid self medication. *NZ Med J* 1986; 99(801): 322-323.

[132] Silvaloganathan K., Johnson P.A., Bray G.P., Williams R. Pericoronitis and accidental Paracetamol overdose. A cautionary tale. *Br Dent J* 1993; 174(2): 69-71.

[133] Bassler K. Use and abuse of high doses of vitamin B6. *Int J Vitamin Nutr Res* 1989; Suppl. 30: 120-126.

[134] Halle A and Sloas D.D. Percutaneous organophosphate poisoning. *South Med J* 1987; 80(9): 1179-1181.

[135] Dicket W.H., Lingeman J.E, Preminger G.M, Smith L.H, Wilson D.M., et al. Laxative abuse as a cause for ammonium urate renal calculi. *J Urol* 1990; 143(2): 244-247.

[136] Simbi K.A., Sechieri S., Rinaldo M., Demi M., Zainardo V. In-utero ductal closure following near-term maternal self medication with numesilide and Acetaminophen. *J Obstet Gynaecol* 2002; 22(4): 440-441.

[137] Oliveira-Fuster G., Mancha-Doblas I., Vazquez-San., McGuel F., Esteva De Antonio I., et al. Surreptitious intake of diuretics as a cause of Pseudo-Barter syndome: apropos of a case and differential diagnosis. *Anales de Medicina Interna* 1996; 13(10): 496-499.

[138] Chem K.C., Meisler D.M., Wilhelmus K.R., Jones D.B., Stern G.A., et al. Corneal anaesthetic abuse and Candida Keratitis. *Opthalmology* 1996; 31(1): 37-40.

[139] Rocha G., Brunnette I., Le-Francois M. Severe toxic keratopathy secondary to topical anaesthetic abuse. *Opthalmology* 1995; 30(4): 198-202.

[140] Ben-Chetrit E and Saint-Jean O. Exacerbation of psoriasis by Ibuprofen. *Cutis* 1986; 38(1): 45.

[141] Villette B., Rybojad M., Puissant A. Dermatologic manifestation of undesirable effects of drugs. *Ann Med Interne Paros* 1989; 140(7): 609-613.

[142] Bengaud S and Saint-Jean O. Properties of drug complications in elderly patients. *Rev Prac* 1990; 40(15): 1366-1370.

[143] Martinez-De-Jesus F.R., Gallardo-Hernandez R., Morales-Guzman M., Peres-Morales A.G. Delay in hospitalization diagnosis and surgical intervention in acute appendicitis. *Rev Gastroenteol Mex* 1995; 60(1): 17-21.

[144] Afolabi A.O, Adekanle O. Nonsteroidal Gastropathy in a Dental Patient: A case report. Nigeria Medical Practitioner 2008; 53(6):110-112.

[145] Lin S.H and Lin M.S. A survey on drug related hospitalization in a community teaching hospital. *Int J Clin Pharmacol Ther Toxicol* 1993; 11(2): 66-69.

[146] Brownlee H.J Jr. Family practitioner's guide to patient self treatment of acute diarrhoea. *Am J Med* 1990; 88(6A): 275-295.

[147] Cornelly C. An emperical study of a model self care in chronic illness. *Clin Nurse Spec* 1993; 7(5): 247-253.

[148] Foster S. Treatment of malaria outside the formal health services. *J Trop Med Hyg* 1995; 98(1): 29-34.

[149] Fleming G.V., Giachello A.L., Andersen R.M, Andrade P. Selfcare, substitute, supplement, or stimulus for formal medical care services. *Med Care* 1984; 22(10): 950-966.

[150] Ruegg A. Contribution of the Pharmacist to safety of self medication. *Soz Praventivmed* 1986; 231(3): 160-164.

[151] Beske F and Hanpft R. Status of self medication in West Germany. *Soz Praventivmed* 1986; 31(3): 156-159.

[152] Meyer U. Thoughts on the qualitative improvement of self medication. *Soz Praventivmed* 1996; 31(3): 166-169.

[153] Viberg N. Selling Drugs or Providing Health Care?: The role of private pharmacies and drugstores, examples from Zimbabwe and Tanzania. *Pharm World Sci* 2009; 29(1): 25-33 (Appendix XLV) Abstract.

[154] Wilkinson I.F., Darby D.N., Mant A. Self care and medication. An evaluation of individual's health care decision. *Med Care* 1987; 25(10): 965-978.

[155] Cranz H. Over-the-counter drugs. The issues. *Drug Saf* 1990; 5 Suppl 1: 120-125.

[156] Mulholland S. Guidelines for management of the problem patient. *Urology* 1988; 32(2 suppl 1): 28-31

[157] Reinstein J.A. OTC Advertising: In whose interest? The Manufacturer's viewpoint. *WHO Drug Information* 1995; 9(1): 11-13.

[158] Kelly J. Implementing a patient's self-medication program. *Rehabil Nurs* 1994; 19(2): 87-90, 95.

[159] Love C.J., Raynor D.K., Coyrtney E.A., Purvis J., Teale C. Effects of self medication programs on knowledge of drugs and compliance with treatment in elderly patients. *Br. Med. J* 1995; 310(6989): 1229-1331.

13

The Relationship Between Alcohol Consumption and Human Immunodeficiency Virus Infection and Risk Behaviour: A Systematic Literature Review of High-Risk Groups, with a Focus on South Africa

Manuela G. Neuman[1,2,*,†], Michelle Schneider[3,†],
Radu M. Nanau[1,2], Charles Parry[3,4] and Matthew Chersich[5,6]
[1]Departments of Pharmacology & Toxicology and International Health, University of Toronto,
[2]In Vitro Drug Safety and Biotechnology, MaRS, Toronto, ON,
[3]Alcohol & Drug Abuse Research Unit, South African Medical Research Council, Cape Town,
[4]Department of Psychiatry, Stellenbosch University, Cape Town,
[5]Centre for Health Policy, School of Public Health University of Witwatersrand,
[6]International Centre for Reproductive Health, Ghent University, Ghent,
[1,2]Canada
[3,4]South Africa
[5,6]Belgium

1. Introduction

The most common mechanism for contracting the human immunodeficiency virus (HIV) is through unprotected sexual intercourse with an infected partner. Biological aspects in the acquisition of HIV such as biological susceptibility to HIV infection are also important, but infection with HIV cannot occur without the behavioural component of bodily fluids mixing between HIV-positive and HIV-negative individuals. This study will focus on the behavioural component of unprotected sexual activity and alcohol consumption in terms of risk for HIV infection.

Physiologically, alcohol use loosens inhibitions and impairs cognitive functions, which may lead to unsafe sexual behaviour, essentially unprotected sex. The relationship between alcohol use and the contracting of HIV is complex, as other confounding variables such as certain personality traits or alcohol expectancies also have an effect on both risky sex and alcohol consumption. These confounding variables may explain or partially explain the observed association between alcohol consumption, unsafe sex and HIV infection.

---

[*] Corresponding Author
[†] Contributed equally to this paper

Alcohol has been implicated in the transmission of HIV via unsafe sex in many systematic reviews examining this association. These literature reviews and meta-analyses have varied foci. Cook and Clark (2005) examined the association between alcohol consumption and sexually transmitted infections (STIs). Baliunas (2010) looked at alcohol consumption and risk for incident HIV. Fischer (2007) and Kalichman (2007) conducted global-level systematic review studies in Africa and sub-Saharan Africa (SSA) respectively. Woolf-King and Maisto (2011) conducted a narrative literature review that included qualitative and quantitative literature on the link of alcohol and high-risk sexual behaviour in sub-Saharan Africa. Pithey and Parry (2009) conducted a descriptive literature review examining studies that quantified the association between alcohol consumption and HIV in SSA. The results included a high-risk group category of shebeen/beerhall patrons and bar and hotel workers. Braithwaite et al., (2007) showed that alcohol consumption impacts negatively on people living with AIDS (PLWA). These reviews and the Braithwaite study, while consistently indicating a strong association, do not provide sufficient epidemiological evidence of causality between alcohol use and HIV sero-conversion. They only demonstrate that alcohol is an important correlate of sexual risk behaviour in populations.

However, event-level studies, such as the diary studies of Leigh (2008) and Room (2008), have provided less-convincing evidence than the aforementioned global-level studies for the link between alcohol use and HIV via unprotected sex. This can be attributed to the fact that event-level studies can better control for confounding variables. Conversely, Kiene (2008) showed in the first diary study in South Africa (SA) that moderate or higher drinking levels prior to sex increased the likelihood of unprotected sex.

Overall, the lack of conclusive evidence of a causal link between alcohol consumption and the transmission of HIV does not justify taking no action (Hill, 1965; Phillips 2004). No action implies that alcohol interventions as HIV prevention or treatment measures are not implemented. In addition, there is evidence that indicates that alcohol consumption has a harmful impact on the pathogenesis of HIV. Alcohol reduces an individual's immune response (Friedman et al., 2006), hence increasing the susceptibility of contracting HIV and other opportunistic infections. For an HIV-positive person on antiretroviral medication, alcohol has a negative effect on HIV treatment adherence (Hendershot et al., 2009). Furthermore, antiretroviral drugs (ARV) and alcohol are both metabolized by the liver. Long-term alcohol abuse can result in liver disease, which in turn affects the ability of the liver to metabolise the ARVs (Shuper et al., 2010). There is sufficient epidemiological evidence that alcohol consumption is linked to HIV progression (Parry et al., 2009). Furthermore, Rehm et al., (2009) and Gmel et al., (2011) modelled the impact of alcohol on ARV adherence and HIV mortality.

In light of the above evidence, the nexus of alcohol misuse and HIV is an important research arena, especially in SA, a country with one of the highest rates of HIV infection globally, as well as one of the highest per drinker alcohol consumption rates (Fritz et al., 2010). Sexually transmitted diseases are the leading risk factor for death and disability in SA, with 98% of these disability adjusted life years (DALYs) due to HIV / AIDS (Johnson et al., 2009). High alcohol consumption rates also translate into an enormous disease burden attributable to alcohol itself (Schneider et al., 2007). It is estimated that 5.6 million people are living with AIDS in SA with the adult prevalence at 17.8% in 2009 (UNAIDS, 2010). As is the case in the rest of SSA, high proportions of the general population in SA abstain from drinking alcohol (Obot 2007; Parry et al. 2005). The per capita annual alcohol consumption in SA is estimated between 10.3 and 12.4 litres (Rehm 2004). However, those who do drink do so to intoxication, particularly during weekends (Parry et al., 2005). In SSA, the alcohol consumption per drinker is 19.5 litres alcohol (Roecke et al., 2008). This figure is slightly higher in SA, at 20 litres per drinker per annum (Rehm et al., 2003). These high alcohol consumption levels for those who drink together with the high background prevalence of HIV do not augur well for the containment of the HIV epidemic in SA.

Alcohol consumption is a multi-dimensional variable that includes volume, consumption patterns, as well as the context of consumption. Although all these factors impact on the outcome of interest, in this case, unsafe sex, the context provides the platform for the physical and psychosocial factors to play out and interact, often synergistically, thus creating greater chances of contracting HIV. Chersich et al., (2010) points out that it is often the context of alcohol consumption that is particularly unsafe for adolescents, sex workers and migrant labourers in SA. These individuals tend to be generally more vulnerable to contracting HIV due to inexperience with alcohol and sex in the case of adolescents, or due to illegal or transient positions for the latter two groups, respectively.

A worrisome trend is the high alcohol consumption in certain sub-populations in SA, particularly emerging adults, farm workers and mine workers. For example, Madu and Matla (2003) reported that 39.1% of emerging adults consume alcohol. Problem drinking in farm workers has been estimated to be in excess of 65% (London, 2000) and over a quarter (26.3%) of HIV-positive mine workers were found to be meeting criteria for alcohol abuse (Säll et al., 2009). Moreover, in a sample of migrant women, alcohol use was a direct predictor of HIV-positive sero-status (Zuma et al., 2003). In a recent South African report, the definition of most at-risk populations (MARPs) for HIV/AIDS was expanded to include persons who drink excessively (Shisana et al., 2009).

Globally, there are two patterns on how the HIV epidemic presents in populations. In many SSA countries, there is a generalised epidemic, as opposed to other countries where it is concentrated in high-risk groups. SA is experiencing a so-called generalised HIV/AIDS epidemic, defined as having an HIV prevalence rate in the general population greater than 1%, with heterosexual intercourse being the predominant mode of HIV transmission. Sexual networking within the population is sufficient to sustain the epidemic, despite high-risk sub-populations that contribute disproportionately to the spread of the disease in SA (Ngom and Clark, 2003).

Notwithstanding the above, in the present review, we discuss the link between alcohol consumption, risky sexual behaviour, and the risk of HIV infection in high risk groups worldwide, with a particular focus on SA. This study will add to the accumulating evidence for delineating the causal pathways of alcohol use and risky sex to HIV infection. Many of the mediating factors in the causal web of alcohol use, unsafe sex and HIV acquisition tend to be more potent and hence easier to ascertain in some of these high risk groups. An important focus for HIV prevention research is to examine the role of alcohol with respect to risky sex in high HIV risk sub-populations.

Universally, sections of populations have higher HIV prevalence rates than the general population. HIV/AIDS MARPs usually include men who have sex with men (MSM), injecting drug users (IDUs) and sex workers. These high-risk groups have larger pooled odds ratios (ORs) of alcohol use and HIV-positive status than the general population (Fischer et al., 2007).

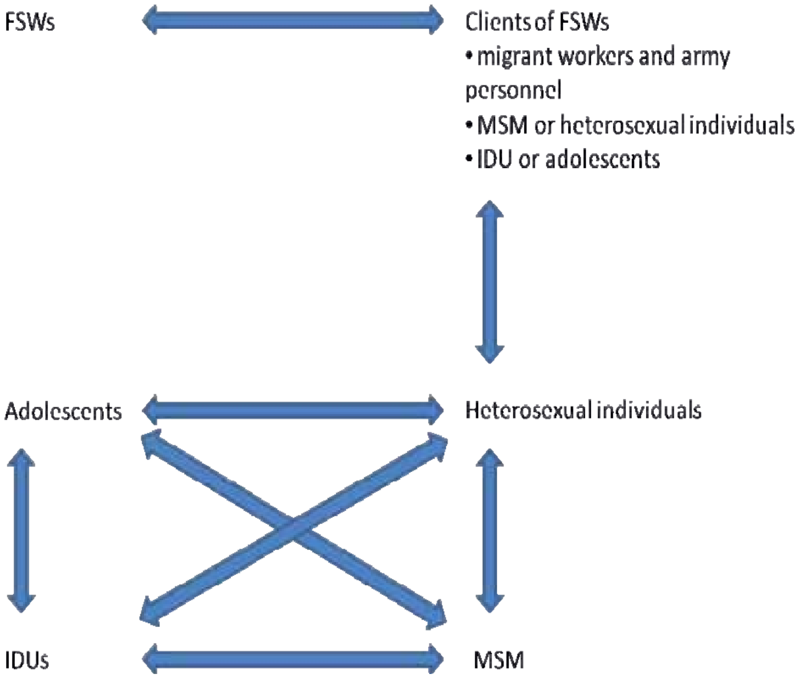


Fig. 1. Depicts female sex workers (FSWs) as key determinants for sex-concurrency for their clients, who include MSM and heterosexual individuals, who can be migrant workers, army personnel, adolescents and/or IDUs.

These high-risk groups engage in patterns of behaviour that carry elevated risk for HIV infection (Parry 2008a). A component of these lifestyles may include closed or interconnected social and sexual networks within these sub-populations with relatively high HIV sero-prevalence rates. In addition, "bridging" effects among individuals from these groups leads to a wider network of HIV transmission. Figure 1 is a model of a network of HIV transmission in the presence or absence of alcohol consumption among the selected high-risk groups reviewed in this paper.

2. Materials and methods

We performed a systematic review of published literature using PubMed, searching for articles that contained information about alcohol drinking patterns and sexual risk behaviours. We limited the search to literature published in English. In order to obtain more focused results so that we could, where necessary, refer to South Africa, we also included the words "South Africa" for the search. However, we did not have "South Africa" as an exclusion criterion. We found over 4000 results using the key words "alcohol" and "HIV", from which we selected 750 of the most recent publications (i.e. June 2008 to June 2010). Reading these articles, the majority were in vitro studies of alcohol or HIV, or were performed on animals, or they described molecular mechanisms or discussed only one of the risk factors; often sexual transmission and alcohol were referred to only once, in passing. For the period January 2008 to October 2010 there were 183 articles that pertained to humans. From the 183 articles, 107articles were selected. The main reason for discarding 76 articles was that they did not discuss alcohol consumption and risky sexual behaviour; the two variables were not linked in any way. The 107 articles were allocated to one or more of the six high-risk groups selected for the study. (Figure 2) Particular attention was paid to those papers that provided an indication of the drinking pattern (e.g. binge drinking, continuous drinking, alcohol abuse and alcohol intoxication) and the amount of alcohol consumed. Where the relationship between alcohol consumption and unprotected sex was quantified in the individual articles, this was captured and reported on.

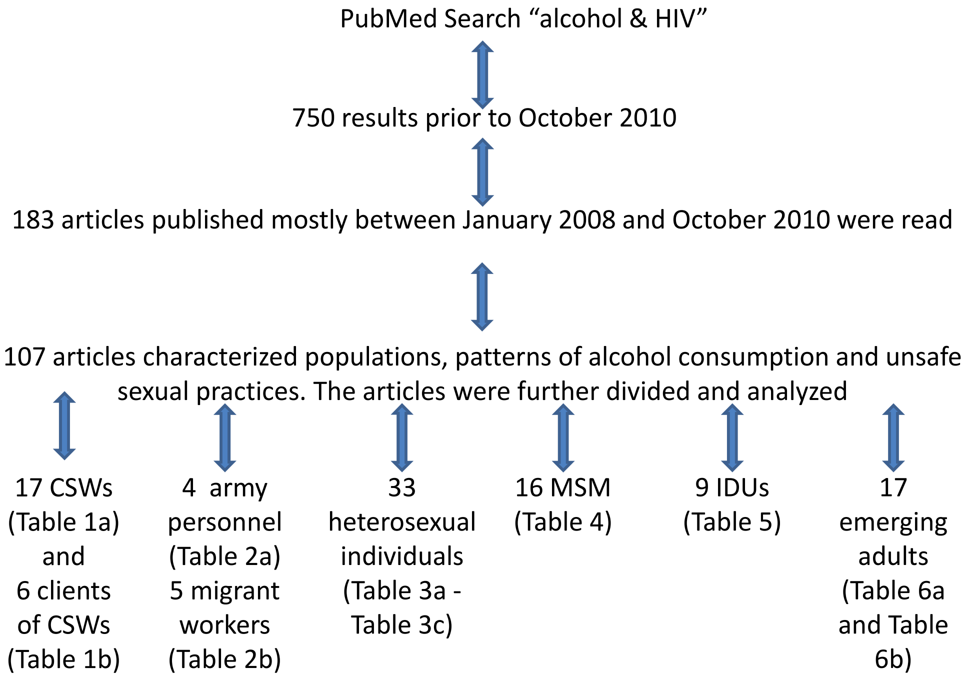


Fig. 2. Flow diagram of the systematic literature review

We also discuss a limited number of relevant papers published prior to 2008, which were used to emphasize some of the points that were not well represented in the more recent publications. These additional papers were obtained via a manual search of the reference sections of the articles obtained from the PubMed search.

3. Results

This section is divided into six subsections, corresponding to each of the six at-risk groups that are discussed in detail.

3.1 Female Sex Workers and their clients

The alcohol and condom use patterns of commercial sex workers (CSW), particularly FSWs, are described in Table 1a. In addition, the alcohol and condom use patterns of clients of the CSWs are described in Table 1b.

Sex with a customer is often accompanied by alcohol consumption, as shown by studies in Afghanistan (Todd et al., 2010), China (Rogers et al., 2002; Wang et al., 2010), India (Bowen et al., 2010; Go et al., 2010; Samet et al., 2010; Verma et al., 2010), Kenya (Chersich et al., 2007), Mongolia (Witte et al., 2010), the Philippines (Chiao et al., 2006; Morisky et al., 2010), Scotland (Thomas et al., 1990), South Africa (Wechsberg et al., 2005; Wechsberg et al., 2009), Tanzania (Fisher et al., 2008; Fisher et al., 2010), the United States (Nemoto et al., 2003), and Vietnam (Nemoto et al., 2008). Alcohol is also consumed by male clients of FSWs in India (Madhivanan et al., 2005; Rodríguez et al., 2010; Samet et al., 2010; Schensul et al., 2010; Verma et al., 2010), Scotland (Morgan Thomas et al., 1990), Sri Lanka (Dissabandara et al., 2009), Thailand (Havanon et al., 1993), and Zimbabwe (Fritz et al., 2002). In addition, FSWs often consume alcohol together with their clients, as observed in massage parlours and bars in Vietnam (Nemoto et al., 2008), the Philippines (Chiao et al., 2006), the United States (Nemoto et al., 2003), China (Wang et al., 2010), India (Rodríguez et al., 2010) and Tanzania (Fisher et al., 2008). An Indian study found that alcohol consumption for both FSWs and their male partners are above the national average for their respective genders (Samet et al., 2010). In a sample of 335 FSWs and 171 female non-SW in South Africa, FSWs were found to be more likely than females who were not sex workers to suffer from an alcohol abuse disorder (Wechsberg et al., 2009). Some women describe alcohol use as a way to please the partner, when this is what the partner wants (Witte et al., 2010).

3.2 Female Sex Workers

A sample of 48 FSWs report that they used alcohol in large amounts in order to decrease inhibitions prior to sex work and to cope with the stigma, psychological distress and violence brought about by their sex life (Witte et al., 2010). Other reasons for alcohol use were to conform to norms of sex partners who preferred them to be intoxicated and would pay more in these situations (Witte et al., 2010). Other groups also report than some FSWs use alcohol to make sex easier (Wechsberg et al., 2005; Nemoto et al., 2008; Todd et al., 2010). Drinking before sex with a client is linked with life-long inconsistent condom use in 454 FSWs (Wang et al., 2010). However, Wang et al. (2010) did not find a direct link between alcohol consumption, condom use, and STD prevalence, supporting the theory by Fisher et al. (2010) that the negative influence of alcohol on condom use is event-specific, rather than global. In support of this, the Mongolian study referred to earlier found that inconsistent condom use also occurs as a result of a partner being willing to pay more for unprotected sex, rather than due to alcohol use per se (Witte et al., 2010).

Generally FSWs who used alcohol before sex with a client, compared with those who did not, were more likely to use condom inconsistently and to be STD-positive (Wang et al., 2010). Eighty five percent of FSWs report that their clients sometimes refused to wear a condom. In a different study in a sample of 3412 FSWs alcohol consumption and HIV prevalence were found to be high whereas condom use was found to be very low suggesting an inverse relationship between alcohol and condom use (Verma et al., 2010).

Alcohol use has a significant association with having a non-paying male partner and having sex with more than three partners per day in 63.7% of FSWs (Verma et al., 2010). Alcohol use prior to sex leads to inconsistent condom use for both male migrant workers as well as for FSWs. HIV/AIDS education and an increase in a sex worker's ability to use a condom effectively have been linked with decreased alcohol consumption and more frequent condom use, particularly among those FSWs who are also IDUs (Fisher et al., 2008; Strathdee et al., 2009; Morisky et al., 2010). A direct correlation between decreasing daily alcohol consumption and increasing consistent condom use was found in a sample of 911 FSWs (Morisky et al., 2010).

Studies conducted in India report inconsistent condom use among FSWs who consume alcohol (Bowen et al., 2010; Go et al., 2010; Rodríguez et al., 2010; Samet et al., 2010; Verma et al., 2010). The rate of inconsistent condom use was found to be very high and no significant differences were found between FSWs who consume alcohol (even in high quantities) and those who do not, in a sample of 211 HIV-positive FSWs (Samet et al., 2010). Fifty seven percent of women report to have used alcohol around the time when they first entered sex work (Bowen et al., 2010). The study in a sample of 220 FSWs (Bowen et al., 2010) found that alcohol and drug consumption has increased since the time these women entered into the sex trade industry.

To complicate matters even further, Chiao et al. (2006) report that alcohol consumption with a customer was accompanied by a significantly higher willingness to use a condom in a sample of 1,114 FSWs. Not surprisingly though, condom failure was also high in intoxicated FSWs or FSWs with intoxicated customers, and STD prevalence was significantly higher in FSWs who had sex with intoxicated customers than in FSWs who did not have sex with intoxicated customers (Chiao et al., 2006). However, alcohol consumption by the FSW did not increase the STD risk any further in this case. Similarly, FSWs who were intoxicated prior to sex were more prone to have STDs than FSWs who were not intoxicated, regardless of the partner's status (Chiao et al., 2006).

In a sample of 159 FSWs (53 working in massage parlours, bars/clubs and the street respectively), massage parlour FSWs report that alcohol consumption was inevitable for them, as it was part of the job (Nemoto et al., 2008). Alcohol consumption was higher for massage parlour and bar/club FSWs, where alcohol was often consumed with customers. In addition, 30% of street-FSWs were also IDUs, using mainly heroin (Nemoto et al., 2008). No significant association was found between condom use and having sex with a customer while under the influence of alcohol. Condom use for vaginal sex with a primary partner was very low, with over 65% of subjects reporting never having used a condom in this situation. Sex under the influence of alcohol was highest for massage parlour FSW and lowest for street FSWs. The trend was the same between sex with casual partners and sex with primary partners (Nemoto et al., 2008).

The risk of condom failure increased significantly if one or both of the partners had been drinking within two hours before sex (Fisher et al., 2010). Condoms are more likely to be used in unfamiliar places, with first-time partners, or if sex was transactional (Fisher et al., 2010). In addition, Fisher et al. (2010) observed that a condom is 10 times more likely to be used if the woman was involved in the decision. In support of this argument, an Indian study found that FSWs avoid using alcohol in order to remain alert throughout the encounter, and to decrease the risk of violence (Rodríguez et al., 2010). This corroborated findings by Go et al. (2010) who report that alcohol consumption before sex, by either the FSWs or her partner, is associated with forced sex in 522 FSWs. In 93 FSWs, Wechsberg et al. (2005) found that FSWs who have been sexually abused are more likely to use condoms inconsistently and to become HIV infected. Daily alcohol consumption was reported by 18% of this sample. Although not significant, women who have been physically abused were more likely to use alcohol daily than those who were never physically abused. In contrast, those women who have never been sexually abused were less likely to consume alcohol daily than those who were sexually abused ($p<0.05$) (Wechsberg et al., 2005). Another study found that binge drinking was statistically associated with sexual violence in 719 FSWs (Chersich et al., 2007).

In a study of Thai army conscripts, several male respondents reported that they use alcohol as a socially-acceptable excuse to not use a condom, and they relied on the FSW to put on the condom (MacQueen et al., 1996). On the same note, Nemoto et al. (2008) explained their finding that condom use was lowest in FSWs working in bars and clubs as customers of these women insisted that a condom is not used. In the same study, one massage parlour FSW argued that she tried to convince her partner to wear a condom, but she sometimes agreed to have unprotected sex because the partner would pay better (Nemoto et al., 2008). Condom negotiation is often hindered by having an inebriated partner. Women often compromise as they have low bargaining power (Nemoto et al., 2008; Rogers et al., 2002). At the same time, fear of violence often diminishes the power of a woman to negotiate condom use (Witte et al., 2010). In comparison, a Chinese study found that 63% of FSWs refused to provide sexual services to partners who refused to wear a condom (Rogers et al., 2002).

Not surprisingly, condom use is not the only determinant of HIV infection. Nemoto et al. (2008) found the highest HIV prevalence (18%) in street FSWs, the group that had the lowest rate of inconsistent condom use. Sex under the influence of alcohol was lowest for street FSWs, with either casual or main partners, and it was highest for massage parlour FSWs. Sex under the influence of alcohol was overall higher with main partner, and condom use was lower (Nemoto et al., 2008). Interestingly, sex under the influence of alcohol, at least with a main partner, was reported by 100% of massage parlour FSWs interviewed, yet none of these women were found to be HIV-positive (Nemoto et al., 2008). The same group also found no HIV-positive Asian masseuses in a sample of 100, although STD prevalence was quite high at 94% (Nemoto et al., 2003). In this case, condom use was higher among Asian-American masseuses, with both casual partners and the main partner (Nemoto et al., 2003). In contrast, a Philippine study reports that sex with an intoxicated customer was higher among street FSWs, and STD prevalence was high in these situations (Chiao et al., 2006).

In India, Verma et al. (2010) found a significant association between alcohol use and having more than three partners per day, both paying and not paying. In these cases, condom use is minimal, and this problem is even more pronounced in migrant FSWs who travel greater distances (Verma et al., 2010). Greater alcohol consumption was associated with a higher number of sex partners in a sample of 1044 FSWs working in bars and hotels (Fisher et al., 2008). It appears to be alcohol consumption, rather than the pattern of drinking, that is linked with HIV risk factors (Chersich et al., 2007). Chersich et al. (2007) found lifetime alcohol consumption, rather than binge drinking, to be linked with being seropositive. At the same time, daily drinking and consuming more than the equivalent of 11 beers in one week are strongly associated with being HIV-positive (Fisher et al., 2008). In addition, HIV risk was positively correlated with the amount of alcohol consumed per drinking occasion, peaking for 3 drinks per occasion, which is less than the level necessary to qualify a woman as a binge drinker. Daily drinkers were nearly four times more likely to be HIV-positive when compared with non-drinkers (Fisher et al., 2008). Alcohol users are more likely to have multiple sex partners (bar and hotel patrons) and are more likely to be HIV-positive (Fisher et al., 2008). In addition, everyday alcohol consumption and binge drinking were significantly associated with inconsistent condom use (Chersich et al., 2007). Despite binging being associated with higher sex risk factors, it did not increase HIV prevalence. Furthermore, binge drinking is associated with sexual violence (OR 1.85, CI 1.27-2.71, p<0.001) and other STDs (OR 1.56, CI 1.00-2.41, p=0.048) (Chersich et al., 2007). Among life-long alcohol abstainers, HIV prevalence was significantly lower than it was among FSWs who had ever consumed alcohol. However, the drinking pattern did not have any significant effect on the HIV status, as the study did not find an association between binge drinking and a higher chance of contracting HIV, than among those who had ever consumed alcohol (Chersich et al., 2007).

3.3 Clients of Commercial Sex Workers

A study of 206 male and three female clients of CSWs found that over 55.0% of CSWs were perceived to be under the influence of alcohol by their clients (Thomas et al., 1990). Interestingly, alcohol consumption did not influence condom use when the CSW was a female, but was negatively associated with condom use with a male CSW (Thomas et al., 1990). Condom use is event-specific for FSWs (Fisher et al., 2010; Wang et al., 2010), and the same can be said for clients of FSWs. Havanon et al. (1993) report that visiting FSWs is a socially acceptable for married men in the Thai society. In a study of 181 male clients of FSWs, condom use was found to be influenced by the perceived “cleanliness” of the establishment and the FSW, as well as the perceived number of sex partner that she has had, rather than one’s state of inebriation. Condom use was higher for students (Havanon et al., 1993). Being drunk is not reported as a reason for not using condoms. In a sample of 84 single male drinkers (aged 18 to 29 years), lower condom use was observed when sex took place away from the brothel, or when a certain level of intimacy with the CSW was achieved. However, no association between alcohol consumption and condom use with a CSW was reported (Schensul et al., 2010).

Sex with a FSW while under the influence of alcohol was high in a sample of 1741 men in an STD clinic. Ninety two percent reported sex with a FSW, with 66% having done so under the influence of alcohol (Madhivanan et al., 2005). Sex while under the influence of alcohol was associated with unprotected sex, anal sex, multiple FSW partners, and a history of STDs (Madhivanan et al., 2005). In a sample of 205 HIV-positive men 26% reported inconsistent condom use with a FSWs and this was found to be correlated with alcohol consumption (Samet et al., 2010). Condom use was slightly higher among younger individuals (Samet et al., 2010), which is similar to findings by Havanon et al. (1993) in Thailand.

In a sample of 324 male drinkers, the number of drinking days was associated with the number of unprotected sex episodes with casual partners of CSWs, as well as having sex while intoxicated (Fritz et al., 2002). At least one episode of sex while intoxicated during the previous six months was reported by 31% of subjects. Sixty nine percent of these men reported doing so with casual partners or CSWs. Having sex while intoxicated was linked with a 20-fold higher chance of having unprotected sex with a casual partner, and a 27-fold higher chance with a CSW. The number of drinking days was correlated with the number of episodes of unprotected sex with casual partners, episodes of paying for sex, and having sex while intoxicated (Fritz et al., 2002). There was also a strong link between having sex while intoxicated and HIV sero-conversion. Of HIV-positive men, 35.7% report to have drunk to intoxication on more than 16 days in the past month. Having sex while intoxicated was associated with a positive HIV status in 37% of the male subjects interviewed. Having sex while intoxicated in the last 6 months was significantly associated with recent sero-conversion (OR 4.5, CI 1.0-19.4) (Fritz et al., 2002).

3.4 Army personnel and migrant workers

The alcohol and condom use patterns of army personnel and migrant workers are described in Table 2a and Table 2b, respectively.

3.4.1 Army personnel

Alcohol abuse is a common problem in army personnel, both active personnel and veterans. Alcohol abuse in this population has been linked with non-monogamous heterosexual sex, including sex with CSWs (MacQueen et al., 1996; Brodine et al., 2003; Tavarez et al., 2010). In a study of 498 sexually-active male military personnel, alcohol abusers were found to be more prone to have multiple sex partners, have sex with CSWs, and use condoms inconsistently. Nineteen percent of individuals, mainly unmarried men, reported having sex with a CSW. The odds of having multiple sex partners were higher in individuals suspected of alcohol abuse. Nineteen percent of the sample reported sex with a CSW, while 93% of this subgroup was believed to suffer from probable alcohol abuse problems. Individuals with suspected alcohol abuse problems were twice as likely to engage in non-monogamous sex, with inconsistent use of condoms. Two hundred and sixteen individuals reported being in non-monogamous relationships, with inconsistent condom use; of these, 86.6% report probable alcohol abuse problems (Tavarez et al., 2010).

In a small sample of 76 young male army conscripts, the majority of the subjects reported using alcohol in a social setting, where all the individuals in the group would drink. Alcoholic beverages were often consumed in brothels, where the subjects would also have access to FSWs. In this context, alcohol was used to decrease inhibitions when interacting with the women, and to increase sexual pleasure (MacQueen et al., 1996). On the other hand, condoms were reported by army personal to decrease sexual pleasure, and alcohol consumption provided a socially acceptable excuse not to use a condom (MacQueen et al., 1996).

Brodine et al. (2003) argued that the types of HIV infections identified in a large cohort of 520 recently-infected HIV-positive military personnel reflected the areas of the world where the soldiers were deployed, providing evidence of unprotected sex, possibly with CSWs. For example, 488 patients were infected with HIV-1 subtype B. Individuals with non-subtype-B HIV were likely to be married, and they were likely to have contacted the virus from CSWs outside the USA. Forty four percent of these subjects were heavy alcohol users, this was higher than the overall HIV cohort (OR 2.3, CI 0.6-10.3) (Brodine et al., 2003).

In 881 HIV-positive veterans, Conigliaro et al. (2003) found that hazardous drinking was common. Hazardous drinking was linked to disease progression, as well as the co-occurrence of other conditions such as hepatic co-morbidity and anaemia. Alarmingly, hazardous drinking predominated in younger individuals and those with detectable viral loads (Conigliaro et al., 2003).

3.5 Migrant workers

As is the case for army personnel, migrant and seasonal workers are another group of individuals who live away from their spouses. Gupta et al. (2010) differentiate between temporarily mobile individuals and permanently mobile individuals. In India for example, most migrants travel in order to provide better conditions for their families, and this may make these individuals less prone to risky sexual behaviour (Gupta et al., 2010). At the same time though, a positive association between the length of time that an individual is mobile for and the number of lifetime sexual partners, including paid sex, was found in a national survey of 124385 women (15 to 49 years) and 74369 men (15 to 54 years) (Gupta et al., 2010). Higher alcohol consumption and a high prevalence of STDs were observed in those migrants who were mobile for the longest periods of time, who used alcohol almost daily, and who engaged in risky sex such as sex with multiple partners and paid sex. A higher incidence of having more than two lifetime partners was found in individuals who used alcohol almost daily (OR 2.94, CI 2.67-3.22, $p<0.001$) (Gupta et al., 2010). Rodríguez et al. (2010) report that male migrant workers consume alcohol in order to build up the courage to seek out FSWs, overcome emotional distress and prolong the sexual encounter. Similar to the study on young army conscripts (MacQueen et al., 1996), the behaviour of inebriated clients mitigates against FSWs negotiating condom use. Fear of violence affected the negotiation of condom use, sometimes derailing it altogether (Rodríguez et al., 2010).

There are also migrant populations that traditionally have low levels of condom use. Xiao et al. (2010) argue that migrants with a low level of education are unlikely to use a condom regardless of whether they use alcohol or not. While no association was found between alcohol consumption of any kind and condom use in migrants, overall alcohol consumption was high in migrant workers (Lin et al., 2005; Rhodes et al., 2010). In addition, a high number of sexual partners, buying or selling sex, and a history of STDs were common occurrences (Lin et al., 2005; Rhodes et al., 2010). In 2153 sexually experienced young rural-to-urban migrants (1425 male and 728 female), levels of intoxication were elevated among migrants compared to the general population (Lin et al., 2005). Intoxication was more prevalent among male migrants than among female migrants ($p<0.001$). Alcohol intoxication was associated with multiple sex partners (OR 3.07, CI 1.91-4.95) and buying sex (OR 5.46, CI 2.97-10.04) in males. Intoxicated respondents were significantly more likely to engage in premarital sex and have multiple sexual partners, as well as purchase and sell sex, compared to non-intoxicated respondents (Lin et al., 2005). In a sample of 100 Mexican migrant workers, 10 (40%) of 25 individuals who report having sex in the past 3 months had done so under the influence of alcohol (Rhodes et al., 2010).

In 7602 male migrants, alcohol consumption was found to be linked to having a higher probability of contacting a FSW, and of engaging in unprotected sex (Verma et al., 2010). Those migrants who contacted FSWs were more likely to use alcohol before sex than other migrants. Both alcohol consumption and unprotected sex were higher for those migrants who were mobile for longer periods of time (Verma et al., 2010). Condom use was furthermore inconsistent when male migrants reported alcohol consumption prior to sex with casual partners (OR 0.7, CI 0.6-0.9, $p<0.01$). Even higher rates of unprotected sex were found when male migrants consumed alcohol and contacted FSWs (OR 2.7, CI 2.1-3.5). About 15% of the total sample of male migrant workers reported sex with FSWs in the last 12 months prior to the survey. The proportion of total male migrant workers who reported sex with both paid (including FSWs) and unpaid partners in the last year prior to the survey was significantly higher among alcohol users than among the non-users. Among the sub-sample of male migrant workers who are clients of FSWs, inconsistent condom use with non-paying casual partners is significantly higher if they used alcohol prior to sex than those who did not consume alcohol prior to sex (Verma et al., 2010). The number of alcoholic drinks per week was positively associated with unprotected intercourse acts with casual partners in the past 3 months (p=0.009) (Amirkhanian et al., 2010).

3.6 Heterosexual couples

The alcohol and condom use patterns of at-risk heterosexual individuals are described in Table 3a, Table 3b and Table 3c.

3.6.1 Heterosexual couples in the United States with unknown HIV status

Raj et al. (2009a) did not find the amount of alcohol consumed before sex to be important with regard to the likelihood of condom use in a sample of 617 at-risk African American men. Fifty three percent of the sample reported never using a condom with the main female partner, and 21% reported never using a condom with casual female partners. Forty nine percent (102) of the sample who reported drinking before sex and 49.1% (53) of the sample who reported drinking to intoxication before sex had unprotected sex. Binge drinking was furthermore found to be associated with sex trade involvement (OR 2.2, CI 1.4-3.5) (Raj et al., 2009a).

In a sample of 301 high-risk males and females, mostly African-Americans living in poor neighbourhoods, binge drinking during the past 30 days was linked with unprotected and casual sex (Towe et al., 2010). However, O'Leary et al. (2006) argued that alcohol consumption is one of several contributors to the HIV epidemic affecting certain at-risk groups, particularly African-Americans.

In a sample analyzing 56 sexual events experienced by 28 homeless women, condoms were used in 19 of these events (Ryan et al., 2009). Condom use was higher when the relationship was perceived as casual, as opposed to more serious. On the other hand, condom use appeared lower when the woman was under the influence of alcohol, however statistical significance could not be established due to the low number of cases investigated (Ryan et al., 2009). Higher condom use with casual partners than main partner was also reported by a sample of 221 incarcerated women (57% identified themselves as white) (Rosengard et al., 2005), as well as 2,864 women (80% African-American) living in high-risk communities throughout the country (Lauby et al., 2001). Binge drinking was reported to lower the intention of the respondents to use a condom at the next sexual encounter with a casual partner (Rosengard et al., 2005). Additionally, the frequency of binging, were negatively correlated with the likelihood of condom use (Lauby et al., 2001). Having exchanged sex for money or drugs was also negatively associated with condom use, particularly with a main partner (Lauby et al., 2001).

In a sample of 136 low-income heterosexual women experiencing physical violence by a male partner (63.2% African-American), Cavanaugh et al. (2010) found that while intoxication did not bring about sexual abuse, it may have influenced the partner's decision not to wear a condom, and 14.0% of the sample report that they were frightened to ask their partner to wear one (Cavanaugh et al., 2010). In a sample of 158 predominantly immigrant Hispanic adult females, the level of intoxication of both the woman and her male sex partner was linked to a younger age of oral sex debut and more life-time sex partners (Dillon et al., 2010). In contrast, Wilson et al. (2010) found that alcohol consumption 5 or more days a week was not linked to having more than 3 sex partners in the past year, but it was associated with all the other risk factors (i.e. having a same sex partner (p=0.01), sex with a CSW (p=0.002), higher prevalence of other STDs (p=0.002)) in a sample of 128 male Mexican immigrants.

At the same time, while binge drinking was correlated with unprotected anal sex in 436 high-risk heterosexual females (70% black), it was only marginally correlated with a past-year history of STDs, but not HIV (Jenness et al., 2011. A history of sexual coercion while under the influence of alcohol was, however, found to be linked with a low likelihood of condom use in 5857 heterosexually active women (67% white). Women who were given alcohol or drugs at coerced sex were more likely to have multiple sex partners and engage in substance abuse (Stockman et al., 2010).

3.6.2 HIV-positive heterosexual couples in the United States

In a sample of 535 HIV-positive African-American couples, alcohol dependency, by either partner, was found to not affect condom use (The NIMH Multisite HIV/STD Prevention Trial for African American Couples Group, 2010). However, this ethnic group is still at risk of HIV infection due to the high number of concurrent sexual partners and concurrent sexual partnerships were found to be especially prevalent among females who scored positive for alcohol abuse (The NIMH Multisite HIV/STD Prevention Trial for African American Couples Group, 2010). While observing that alcohol use decreased the frequency of condom use in 326 AIDS patients, Gerbi et al. (2009) found that ethnicity does not influence this behavior. The frequency of alcohol use was correlated with a higher number of sex partners (p=0.003) and lower condom use (p=0.001) (Gerbi et al., 2009).

In a sample of 187 sexually active HIV-positive women (aged 18–61) in ambulatory care, binge drinking was found to double the likelihood that a condom was not used at last vaginal sex. Twenty five percent of the sample was classified as binge drinkers (Theall et al., 2007). Sixty two percent of women were found to have used condoms inconsistently. Although alcohol consumption had no influence on the woman's preference to use a condom, the partner has an easier time manipulating the woman into engaging in unprotected sex when the woman was under the influence of alcohol (Theall et al., 2007). In a predominantly white sample of 262 patients (i.e. only 23.7% of participants were African-American) alcohol use is still associated with a higher likelihood of having sex with multiple partners (Stein et al., 2005). In this sample, the negative effect of alcohol on condom use becomes more apparent. Both the likelihood of having any sex and that that sex was unprotected were associated with any alcohol use, number of alcohol use days, number of drinks per drinking day, number of binge drinking days, and hazardous drinking. (Stein et al., 2005).

3.6.3 Heterosexual couples outside of the United States

In a sample of 1370 women in Tanzania, Kapiga et al. (1998) observed that 5.5% of their sample (3.8% of non-drinkers and 9.4% of drinkers) seroconverted during the period between baseline and follow-up (i.e. a period of 1 to 3 years). Even though condom use was comparable between alcohol drinkers and non-drinkers, sero-conversion risk was found to be higher for those who consumed alcohol (Kapiga et al., 1998). In a Tanzanian sample of men, Ghebremichael and Paintsil (2009) found condom use with a main partner to be low after alcohol consumption in a sample of 789 men, even though 6.5% of the subjects were identified to be HIV-positive. Alcohol abuse was associated with higher STD prevalence. Most of these individuals (88%) were in monogamous relationships, and condom use was therefore low. The only risk behaviour associated with HIV was numerous sex partners in the past 3 years (12% of population) (Ghebremichael and Paintsil, 2009).

Alcohol may enable males to release repressed feelings brought about by social hardship (Emusu et al., 2009), and alcohol abuse by the male partner is a strong indicator of both physical and sexual violence in African countries, with rape being common in these situations (Phorano et al., 2005; Seedat et al., 2009). Since sex is often unprotected in these cases, both the perpetrator and the victim are in danger of acquiring HIV in the case of sero-discordant partnerships. Alcohol abuse by the male partners of 26 women was associated with sexual violence and the sexual abuse of women in this study (Emusu et al., 2009).

In South Africa, sexual violence brought about by alcohol abuse was identified in a sample of 428 men with multiple concurrent female sexual partners. Again, condom use was found to be low in these episodes of sexual violence (Townsend et al., 2011). Alcohol consumption (55%, CI 49.3%-60.2%) and unprotected sex (76.5%, CI 71.5%-81.3%) were found to be HIV risk factors (Townsend et al., 2011).

In a sample of 292 men and 219 women in STD care in South Africa, individuals with a drinking problem were also more likely than individuals without a drinking problem to also be IDUs and to share injection equipment (OR 6.3, CI 2.3 to 17.2, $p<0.01$), have had an IDU partner (OR 4.6, CI 2.1 to 10.2, $p<0.01$), have two or more sex partners in the past 3 months (OR 3.0, CI 1.9 to 4.4, $p<0.01$) and to exchange sex for money or a place to stay (OR 4.8, CI 2.4 to 9.2, $p<0.01$) (Kalichman et al., 2006). However, no association between problem drinking and engaging in MSM sex was found (Kalichman et al., 2006). Unprotected serodiscordant sex was identified in 1052 men and 679 women being treated in an STD clinic. While the likelihood of using a condom for each serodiscordant sex episode was high overall, alcohol consumption was significantly associated with an increased number of sexual partners, thus leading to a relatively high number of unprotected incidents (Kalichman et al., 2010). Alcohol use before sex was associated with HIV-positive individuals (12 of 34, 36%) and engaging in unprotected serodiscordant sex versus protected sex (31 of 184, 26%) (OR 2.1, CI 0.9-5.0, $p>0.05$) (Kalichman et al., 2010).

In a sample of 2,618 primary care patients in South Africa (63.8% female and 36.2% male), the prevalence of risky sexual behavior was 26% (Avalos et al., 2010). Sexual risk taking was overall higher for younger individuals (18 to 24 years old), and for men. Among people reporting at least one sexual risk behavior, 51.9% reported hazardous alcohol consumption ($p<0.001$). Hazardous alcohol consumption was associated with five of the six sexual risk behaviours analyzed: having a partner who ever traded sex for drugs, transportation, or money; having a partner who used injection drugs; having a partner who had an STI; having multiple partners; or failing to use a condom at last intercourse ($p<0.05$ for all risk behaviours) (Avalos et al., 2010). Similar findings emerged from a study conducted on 488 participants (Andersson et al., 2009). Again, male gender was significantly associated with more sex partners in the previous six months ($p<0.001$), more casual/anonymous partners ($p<0.001$) and more one-night stands ($p<0.001$). Perhaps most worrisome is the finding that males engaged in unprotected sex with known/suspected HIV-positive partners (Andersson et al., 2009).

In another South African study of 395 participants (195 males and 200 females), Wong et al. (2008) observed that while alcohol abuse is still higher in men than in women ($p=0.001$), women who were recently abused (past 30 days) by their partner were more likely than not to suffer from problem drinking (OR 3.0, CI 1.5 – 5.9, $p=0.0005$) and depression (odds ratio 3.1, CI 1.5 – 6.2, $p=0.005$). No correlation between intimate partner violence and depression, alcohol abuse, and sexual risk behaviours were identified in men. Any 30 day alcohol use was only marginally associated with intimate partner violence ($p=0.08$). For men, intimate partner violence was however linked to the abuse of various drugs ($p=0.02$) (Wong et al., 2008). Lifetime and past 6 month exposure to intimate partner violence were both comparable between males and females. Lifetime exposure to intimate partner violence was very close to 100% for both genders. Also, rates of depression and sexual risk behaviour were comparable between men and women (Wong et al., 2008).

In a sample of 112 women [60 Black and 52 Coloured (mixed race) persons], ethnic differences in terms of drinking patterns were observed, despite both groups reporting that they had sex while under the influence of alcohol (Wechsberg et al., 2008). For Black women, more alcohol consumption was linked to more frequent sex, although they often reported having only one partner. On the other hand, Coloured women were more likely to have multiple partners (1.53 partners in the past 30 days for Coloured women, vs. 0.98 for Black women) (Wechsberg et al., 2008).

Hargreaves et al. (2002) found an interesting link between alcohol use, age, socio-economic status and HIV prevalence. In a sample of 622 males and 893 females, they observed that alcohol consumption increased with both age and socio-economic status, for both men and women (Hargreaves et al., 2002). Drinking alcohol in the last month was significantly associated with HIV infection for both males and females aged 25–49 years. For both males (OR 1.7, CI 1.0-2.8) and females (OR 1.8, CI 1.0-3.3), there was a high correlation between high alcohol consumption and HIV acquisition. There is also a correlation between non-condom use and HIV infection (OR 3.1, CI 1.3-7.4). Condom use was higher for younger individuals in general, and for women with a higher socio-economic status. Overall, Hargreaves et al., (2002) found that past month drinking was significantly associated with HIV infection in individuals 25-49 years of age.

In a sample of 181 alcohol and/or drug-dependent men and women in India, binge drinking was found to be negatively associated with being HIV-positive, or also being an IDU (Raj et al., 2009b). Moreover 40% of the sample was also IDUs, and 70% of the entire sample reported two or more sex partners in the past three months. Raj et al. (2009b) argue that the effects of alcohol on HIV infection may be masked by the already low rate of condom use in their sample.

In another study conducted in India, current alcohol consumption was found to be associated with higher odds of having premarital sex in a sample of 1,642 never married males and 778 never married females aged 15–24 years. This association was significant in males only (Kumar et al., 2010).

Being under the influence of alcohol by the man was also reported as a reason for risky premarital and extramarital sex among samples of both HIV-positive (Thomas et al., 2009) and HIV-negative individuals in India (Berg et al., 2010). Condom use was very infrequent among males in the HIV-positive sample, comprising of 100 women and 77 men, and it was nonexistent among women. The women however did not blame alcohol consumption for their risky behaviour (Thomas et al., 2009). In an HIV-negative sample of 486 married men living with their wives, drinking was found to lead to domestic violence (Berg et al., 2010), similar to some of the African studies reviewed above. Drinking in public venues means that these men would also have access to FSWs, thus increasing both their and their spouses' risk of HIV infection (Berg et al., 2010). In addition, a direct correlation between the level of alcohol consumption and the degree of domestic violence was found. Heavy drinkers (two or three times a week or more and have three or four drinks or more on a typical day when they are drinking) were 3.5 times and 6.5 times more likely to engage in this behaviour when compared to overindulgent (once a week or less and have three or four drinks or more on a typical day when they are drinking) drinkers and social drinkers (one or two drinks on days in which they drink), respectively (Berg et al., 2010).

In a large sample of 12,617 subjects in India, alcohol consumption was linked to new HIV infections in men only (Dandona et al., 2008). For women, the only significant HIV risk factor was multiple male sex partners (OR 17.85, CI 4.20-75.84) (Dandona et al., 2008). The severity of the alcohol misuse shows a linear association with HIV risk taking in a sample 1137 males (Nayak et al., 2010). Compared to non-hazardous drinking, alcohol abuse (OR 2.35, $p<0.05$) and alcohol dependence (OR 3.55, $p<0.001$) were significantly associated with risk behaviour. Among 433 drinkers (38.1% of the entire sample), the prevalence of hazardous drinking was 56.4%. While condom use is not discussed, current drinking is associated with HIV risk factors in general (OR 6.15, CI 3.70-10.22, $p<0.001$) (Nayak et al., 2010). S.K. Singh et al., (2010) in a study conducted in India, found that married men find themselves in situations where they are expected to drink more and to have sex with partners other than their spouses, but at the same time they are expected to use condoms with these partners. As a consequence, condom use with casual partners was found to be 14 times higher than with main partners, regardless of whether the men were married or not ($p<0.01$) (S.K. Singh et al., 2010).

3.7 Men Who Have Sex with Men

Using a large sample of 1,050 in-care HIV-positive individuals (496 MSM (47% of the entire sample)), Golin et al. (2009) found that binge drinking at least once a week and alcohol consumption before sex were more prevalent in MSM than they were in women and heterosexual men (Golin et al., 2009). The likelihood of having unprotected sex was higher for MSM than it was for men who have sex with females (MSW), but it was lower than it was for women. In addition, the proportion of MSM who reported that alcohol consumption made sex less safe was higher than the sample average (Golin et al., 2009). Alcohol, at all levels of use, was associated with increased sexual risk taking in a sample of 262 HIV-positive individuals (47 MSM). In a study of 166 sexually active individuals, 99 (59.6%) reported unprotected sex (Stein et al., 2005). Increased odds of having any form of sex, including unprotected sex, was associated with any use of alcohol, number of alcohol use days, number of drinks per drinking day, number of binge drinking days, and hazardous drinking. In this sample, MSM were identified as engaging in unprotected sex almost three times as often as any other group, including IDUs ($p<0.05$) (Stein et al., 2005). In general, hazardous drinkers were found to be 5.64 times more likely to have multiple partners and to engage in unprotected sex, when compared to non-hazardous drinkers ($p<0.01$). This was once again more pronounced in MSM (Stein et al., 2005). The number of drinks on a typical drinking day was positively associated with unprotected sex in 321 methamphetamine-using, HIV-positive MSM ($p<0.05$) (Semple et al., 2010). In a sample of 478 AIDS-positive MSM, Bouhnik et al., (2007) found that unprotected sex with casual partners was widespread, and it was even more frequent with main partners than it is with casual partners, putting the regular partner at risk of HIV-infection. In comparison to their American counterparts, French MSM are more prone to unprotected sero-discordant sex following binge drinking (Bouhnik et al., 2007). While protected sex is higher with casual partners than with regular partners, further evidence that alcohol consumption still lowers the likelihood of protected sex with a casual partner is provided by Folch et al. (2009). In a cohort of 850 MSM, alcohol use before sex was associated with unprotected sex with casual partners (Folch et al., 2009).

Alcohol use before sex has been directly linked with HIV sero-conversion. In a large cohort of 4,295 initially HIV-negative MSM who were in a non-monogamous relationship with an HIV-negative partner, Koblin et al. (2006) attribute 29% of sero-conversions within a 48 month period to alcohol use. Overall, 72.1% of men reported using alcohol or drugs before having sex. Sero-conversion was achieved mainly through unprotected sex with a large number of sex partners. The highest risk of sero-conversion (32.3%) was associated with having four or more male partners (OR 2.84, CI 1.72 to 2.69) (Koblin et al., 2006). Problem drinking was associated with unprotected sex with a sero-discordant male partner, as well as unprotected vaginal or anal sex with female partners and transgender partners among 197 African-American MSM (Reisner et al., 2010). Problem drinking was also associated with unprotected sex with a transgender person (OR 5.23, CI 1.26-21.69, $p<0.02$) and unprotected vaginal or anal sex with a female (OR 3.25, CI 1.70-6.24, $p<0.004$) (Reisner et al., 2010).

The link between education and alcohol abuse in MSM populations is unclear. While Reisner et al., (2010) in the US and Tripathi et al., (2009) in Estonia in a sample of 79 MSM report that alcohol being linked to unprotected sex is a more frequent problem among individuals without a university degree, Mackesy-Amiti et al., (2010) found alcohol dependence to be relatively high in a sample of 187 MSM made up predominantly of employed individuals with a college education or higher. Even in this sample of well-educated individuals, both receptive and insertive unprotected anal sex were reported by almost one third of the sample, and 32% reported sex with a partner whose HIV status was positive or unknown (Mackesy-Amiti et al., 2010). In a sample of 378 black MSM, higher monthly income, as well as purchasing and exchanging alcohol and drugs for sex were linked with a higher likelihood of being HIV-positive (Lane et al., 2009).

The risky sexual behaviour of MSM is likely to have wider ramifications than the group itself, as 83.8% of a sample of 68 African-American MSM report concurrent sexual relations with both males and females (Operario et al., 2009). Alcohol consumption and sex while under the influence of alcohol were high in this population, as were concurrent unprotected relationships with both males and females. There was evidence for concurrent partnerships across gender groups as well as concurrent unprotected sex partners. Overall, 73.5% reported having had recent (3 month) concurrent sexual partnerships with more than one gender group (Operario et al., 2009).

Alcohol consumption has been found to be associated with unprotected MSM in other countries as well. Studies from Canada (Lambert et al., 2009), Australia (Prestage et al., 2009), Spain (Folch et al., 2009), Estonia (Tripathi et al., 2009) and Mexico (Mendoza-Pérez et al., 2009) report this. In addition, Lambert et al., (2009) report alcohol consumption two hours before sex was higher when the partner was casual, than when the partner was in a stable relationship with the subject of the study in a sample of 965 MSM who reported having sex with a partner with whom they were not in a couple relationship at last sexual episode. Van Griensven et al., (2010) found that alcohol consumption led 823 MSM to have more frequent sex, as well as more unplanned sex, with both casual and male CSWs. At the same time, sex frequency was not linked to condom use (van Griensven et al., 2010). In a sample of 566 MSM, Tsui and Lau (2010) argue that the way in which an MSM picks his partner will determine the type of risk that he is willing to take, in terms of condom use, as well as the likelihood that they will consume alcohol prior to sex. To this avail, Chinese MSM who recruit their partners from public venues are more likely to consume alcohol and to have multiple sex partners, while MSM who recruit their partners through the internet are more likely to have unprotected sex, be infected with STDs, and buy or sell sex, independent of alcohol consumption (Tsui and Lau, 2010).

3.8 Injecting drug users

Being an injecting drug user is a risk factor for HIV infection (Sander et al., 2010). In a large, longitudinal study of drug users (72% male, 90% African-American) IDUs were found to be at increased risk of greater alcohol consumption compared to non-IDUs. As greater alcohol consumption is further linked to having more sex partners, a greater risk of HIV infection exists among IDUs (Sander et al., 2010). Among IDUs, alcohol consumption and binge drinking in particular have been found to be associated with needle sharing (Arasteh and Des Jarlais, 2009; Matos et al., 2004; Poudel et al., 2010; Stein et al., 2000), multiple sex partners (Arasteh and Des Jarlais, 2009 Chan et al., 2010 Matos et al., 2004; Poudel et al., 2010), sex under the influence of alcohol and/or drugs (Chan et al., 2010), sex with CSWs (Arasteh and Des Jarlais, 2009), sex with a paying customer (Matos et al., 2004), unprotected sex (Arasteh and Des Jarlais, 2009; Chan et al., 2010; Jenness et al., 2010; Matos et al., 2004), and injecting three or more times a day (Matos et al., 2004). For example, the odds ratio for sharing needles while intoxicated (compared to being sober) was 2.1 (CI 1.1-4.3) in a sample of 557 IDUs (89.4% male) (Matos et al., 2004). In the same sample, alcohol intoxication is associated with exchanging sex for money or drugs (Matos et al., 2004). Increasing alcohol use and alcohol addiction was associated with more frequent needle sharing and increased HIV transmission in a population of 196 (68% male, 85% white) active IDUs in care (Stein et al., 2000). In this sample, alcohol abusers were more likely to share needles than non-abusers (OR 2.3, CI 1.2–4.4, p=0.01), and a direct correlation between increasing alcohol consumption and more frequent needle sharing was found (Stein et al., 2000). Condom use was higher when neither partner was intoxicated, and it was higher overall when the partner was deemed as casual, as opposed to a main partner in 1253 HIV-positive IDUs (81% male, 50% Hispanic, 36% African-American) (Arasteh and Des Jarlais, 2009).

Using a large sample of 9,519 adolescent IDUs (71.6% male, 41.4% Caucasian), Chan et al. (2010) found that the degree of risk taking is comparable between male and female subjects. The most prevalent sexual risk behaviour was having multiple sexual partners (39.3% of the sample). Gender differences were still identified, with women reporting more unprotected sex than men as well as sex with more IDU partners, while men reported more sex partners than women (Chan et al., 2010). Moreover, women were more likely to be IDUs themselves, when compared with men. These women were also more likely to have sex while under the influence of alcohol or drugs, and to trade sex for money or drugs (Chan et al., 2010). About 7% of adolescents reported using alcohol or drugs to make sex last longer or hurt less. Older age was associated with higher risk taking in terms of sex while under the influence of alcohol or drugs, unprotected sex, and multiple sex partners (Chan et al., 2010). On the other hand, higher education, older age and female gender were both associated with lower alcohol consumption, particularly binge drinking, among IDUs (Sander et al., 2010). African-American youth were less likely to engage in sexual activities after consuming alcohol or drugs, yet they were generally more likely to engage in sex with multiple partners. Additionally, there seems to be an association between the severity of substance abuse and the degree of risk taking leading to STDs (Chan et al., 2010).

The overall risk of having an IDU partner among 601 at-risk, non-IDU heterosexuals (57.4% female, 78.5 Black) was 13.8%. Binge drinking at least once a week was associated with a significantly higher risk of having an IDU sexual partner (OR 1.73, CI 1.08–2.76, p=0.02) (Jenness et al., 2010). Partners of IDUs were found to abuse alcohol and non-injectable drugs, and to practice unprotected sex with multiple partners (Jenness et al., 2010). HIV prevalence was high in this cohort of heterosexual New York individuals, and higher odds of testing positive for HIV were found among older individuals, and among those who had IDU sex partners (Jenness et al., 2010). A high prevalence of HIV was reported in another study analyzing a cohort of 296 IDUs (Poudel et al., 2010). In this study, 59% of 213 sexually-active participants reported multiple sex partners (Poudel et al., 2010).

In a sample of 240 IDUs, Parry et al. (2008a) found that needle sharing takes place primarily with close friends and less frequently with strangers. In addition, HIV-positive IDUs report being more responsible about their disposal practices (Parry et al., 2008a). Interestingly, Parry et al., (2008a) found that IDUs often have sex with CSWs and MSM, thus creating a "bridging" effect, where HIV can be spread among individuals from these three vulnerable groups. At the same time, IDUs can also be CSWs or MSM themselves. In a sample of 78 IDU MSM, drug use led to sexual risk taking and needle sharing, despite HIV risk knowledge being high (Parry et al., 2008a). Further "bridging" was observed in this study, as these MSM would have sex when high with both men and women (Parry et al., 2008a).

3.9 Emerging adults

3.9.1 United States

The findings of studies looking at the link between alcohol and HIV risk behaviours are listed in Table xx. When looking at emerging adults in the United States, males seem to be the ones more likely to consume alcohol prior to sex (Apostolopoulos et al., 2003; Murphy et al., 2009, Alleyne et al., 2010; Nkansah-Amankra et al., 2010). The major risk factor associated with alcohol consumption was identified as having multiple sex partners

(Alleyne et al., 2010; Nkansah-Amankra et al., 2010). Within a sample of 1,474 high school students (grades 9–12), Latino males were found to take the most risks in terms of sexual behavior (Nkansah-Amankra et al., 2010). These risks include low HIV education, high alcohol consumption, particularly current binge drinking, and a high number of sex partners (Nkansah-Amankra et al., 2010). Ethnicity has also been linked to risk taking for university students. For example, in a sample of 222 black students and 335 white students, Hou (2009) found that African-American students were safer in terms of condom use and alcohol consumption prior to sex, but at the same time were 1.71 times as likely to engage in vaginal sex as white students, and they start doing so at a younger age (Hou, 2009). Another study looking at a large sample of American adolescents also found a lower level of sexual abstinence among African-American subjects, but at the same time more individuals in this group indicated that they always or almost always used a condom. Values for white and Latino individuals were comparable (Murphy et al., 2009). Alcohol consumption was found to be lowest in African-American women within a sample of 425 undergraduate-students enrolled full-time (Randolph et al., 2009). Frequent drinking for non-African-American women and frequent binge drinking for older men were once again linked to a higher number of sex partners. At the same time, younger age and a better understanding of the HIV risk were associated with higher prevalence of condom use for both men and women (Adefuye et al., 2009; Randolph et al., 2009). In addition, certain personality traits, for example sensation seeking, impulsivity, and disinhibited behavvior due to alcohol consumption were indicators of unprotected sex in two samples (n=270 and n=490, respectively) of sexually-active college students (Sheer and Cline, 1995; Xiao et al., 2010). In a sample of 390 students, having consumed at least one drink in the past 30 days was associated with not using a condom in both males (OR 1.24, CI 0.57-2.72, $p \leq 0.05$) and females (OR 1.81, CI 1.06-3.10, $p=0.04$). Having consumed alcohol in the past 30 days was a predictor of no condom use for females, especially those over 30 years of age (OR 3.43, CI 1.33-8.86, $p=0.01$) (Adefuye et al., 2009). The implications of this finding must be considered in light of the fact that many of these women also report fewer partners, and condom use is overall low in monogamous relationships (Adefuye et al., 2009). No link between number of sexual partners and condom use was found (Sheer and Cline, 1995; Randolph et al., 2009). At the same time, planned sex (OR 1.28, CI 1.04–1.59) and sex with a casual partner (OR 3.84, CI 2.30-6.41) were linked with higher condom prevalence in a population of 112 adolescents (Morrison et al., 2003). While Morrison et al. (2003) did not find a link between either alcohol consumption or the amount of alcohol consumed before sex and condom use, Murphy et al. (2009) found a direct positive correlation between the amount of alcohol consumed and the degree of risk taking in a large sample of 8,208 youth. Unfortunately, the latter study failed to separate condom use from the number of sexual partners, when discussing high risk behaviour (Murphy et al., 2009).

Two other at-risk categories of emerging adults have been identified: club goers and spring breakers. Binge drinking was associated with sex after drinking in a sample of 308 young adults at nightclubs (Wells et al., 2010). The number of drinking days, especially binge drinking days, was positively associated with sex after drinking, and white subjects indicated more frequent drinking. Younger club-goers were more likely to have sex after drinking than older respondents (OR 1.75, CI 1.01-3.03). Younger club-goers also reported less safe sex after drinking (OR 2.34, CI 1.22-4.50) (Wells et al., 2010). For this group, drinking frequency was associated with less safe sex, however the amount of alcohol consumed was not important, as no association between unsafe sex and binge drinking was found (Wells et al., 2010). In a sample of 532 spring breakers (321 female and 211 male), 49% of men and 38% of women reported having sex as a direct result of drinking (Apostolopoulos et al., 2003). In this group, one third of individuals report that alcohol consumption led them to have unprotected sex, and the ratios were comparable between males and females (Apostolopoulos et al., 2003).

Alcohol drinking has been associated with dating violence in a population of 2438 high school students (grades 9 to 12) (Alleyne et al., 2010). Alcohol consumption during the last sexual experience was significantly higher in males than in females ($p<0.01$). At the same time, condom use ($p<0.001$) and multiple sex partners in the last 3 months ($p<0.01$) were also higher in males. Interestingly, men reported to have experienced more dating violence, but females have experienced more forced sex (Alleyne et al., 2010). In a study on undergraduate students in southern US, older respondents reported more frequent binge drinking episodes, but these are once again accompanied by a higher number of sex partners and a higher likelihood of condom use (Randolph et al., 2009). The group reporting the lowest alcohol use was African-American females (Randolph et al., 2009).

3.9.2 Outside of the United States

Young individuals use alcohol "to have fun". A study conducted among 490 sexually active college students in China revealed that despite knowing the risk associated with unprotected sex, alcohol consumption diminished condom use (Xiao et al., 2010). While there seems to be almost general consensus that alcohol consumption leads to sex, the link between alcohol use and condom use is less clear. For example, alcohol use was linked to sexual initiation in two samples of secondary school students. However, condom use was very high in a sample of 768 students (Tavares et al., 2009), while it was more inconsistent in a sample of 3,575 students (Campo-Arias et al., 2010). In the latter study, alcohol consumption was linked to risky sexual behaviour (OR 2.50, CI 95%, 1.3-5.1) (Campo-Arias et al., 2010). Tavares et al. (2009) attribute these differences to a better education and wider access to information regarding HIV/AIDS that was available to the students in their sample.

Several African studies, particularly from South Africa, will be discussed below. In a sample of 511 individuals, Singh K. et al. (2010) found that the highest number of sexually-active teenagers was among those recruited from venues that served alcohol. Individuals recruited in the nightlife/drinking venues, particularly 15 to 24 year olds reported the greatest alcohol consumption. Among them particularly women with the highest number of sexual partners reported the riskiest sex. This sample does not report the highest condom use (Singh K. et al., 2010). In a large sample of 4,724 young women and 4029 young men (12–25 years), having ever used alcohol was significantly associated with a lower age of first sex for both men and women (McGrath et al., 2009). In a study on 661 grade 9 students, alcohol and HIV prevention education affected HIV-related risk factors (Karnell et al., 2006). While it did not have any effects on alcohol-related risk factors in those individuals who had already had sex before pre-test, fewer students who had not had sex prior to pre-test were likely to drink or to report that their partner drank prior to sex. Sexual activities while under the influence of alcohol were significantly reduced ($p<0.05$), and females increased their sex refusal self-efficacy ($p<0.05$) following intervention (Karnell et al., 2006). Perhaps worrisome is the finding by Morojele et al. (2006b) that some of the boys from a sample of 61 adolescents (12 to 17 years) stated that they could obtain a more positive status by having multiple sex partners, and by having unprotected sex. In addition, some boys enjoyed the thrill of having multiple sex partners (Morojele et al., 2006b).

4. Discussion

It is difficult to summarise the findings across the studies within the selected high risk groups and more so across the high risk groups because of the varied study designs, the difficulty of accurate measurements of the variables and the complexity of the subject matter. Female commercial sex workers are a particularly vulnerable group for contracting HIV group as they experience additional risk for contracting HIV resulting from biological susceptibility and factors related to gender inequality that often involves economic dependence on sexual partners. However this designation, namely the most at risk of the high risk groups may not be that relevant as there are often, sexual partnerships across the high risk groups as well as between the high risk groups and the general population.

4.1 Female Sex Workers (Table 1a)

The 17 selected studies in this group reflect a wide range of alcohol use patterns. HIV status is not reported in many of the studies. In general, alcohol use is more likely lead to inconsistent condom use. Chersich et al. (2007) in a study conducted in Mombasa, Kenya found that binge drinking was associated with inconsistent condom use (OR 1.59, CI 1.00-2.53, p=0.047). Furthermore, condom use appears to be a function of the situation in which a woman finds herself in, and the amount of control that she has over the situation. The use of alcohol by commercial sex workers and their clients also leads to more aggressive behaviour and sexual violence; the latter generally excludes condom use.

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status
Todd et al. 2010	Jalalabad, Kabul, and Mazar-I-Sharif, Afghanistan	Only 4.7% of 520 FSWs report having ever used alcohol	30.2% of FSWs had ever used a condom with a client Of these, 38.2% (60) report always using condoms with clients	Prevalence of HIV was 0.19%
Rogers et al. 2002	Beijing, China	42% of women and 32% of their clients report alcohol consumption during sex work	61% of women reported consistent condom use	
Wang et al. 2010	Nanning, Guangxi Zhuang, China	29.4% of women reported having had sex with their clients after drinking alcohol	Inconsistent condom use over their life time was significantly associated with drinking alcohol before having sex with a client ($p<0.05$)	
Bowen et al. 2010	Nagaland, India	Alcohol use is widespread 68.5% of FSWs interviewed were regularly using alcohol or other drugs	Condom were used 65.3% of times in the past week	
Go et al. 2010	Chennai, Tamil Nadu, India	Women who had a strong tendency to drink alcohol before sex were more likely to have more partners and to have experienced forced sex	Women who reported >20 days of alcohol consumption in the last 30 days were more likely to have unprotected sex	

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status
Samet et al. 2010	Mumbai, India	Overall, 38% of FSWs drink 11% are alcohol dependent 32% are heavy alcohol consumers (>7 drinks per week or >3 drinks on a single occasion)	90% of women reported inconsistent condom use during transactional sex over the last year	HIV-positive
Verma et al. 2010	India	62.0% report alcohol consumption in the last month 53.8% report alcohol consumption prior to sex	Overall, inconsistent condom use was reported 58.5% of times	
Chersich et al., 2007	Mombasa, Kenya	33.0% were binge drinker 44.7% were non-binge drinkers 22.4% were abstainers	Binge drinking was associated with inconsistent condom use (OR 1.59, CI 1.00-2.53, p=0.047)	39.9% who had ever consumed alcohol were HIV-positive 37.3% (84/225) of binge drinkers 23.2% (36/155) of lifetime-alcohol abstainers
Witte et al. 2010	Ulaanbaatar and Darkhan Uul, Mongolia	92% of women consume alcohol on a typical day (19%, 1 to 2 drinks; 29%, 3 to 4 drinks; 27%, 5 to 6 drinks; 4%, 7 to 9 drinks; 13%, 10 or more drinks) 44% (n=21) consumed five or more drinks per day	69% (n=33) reported using condoms inconsistently with paying partners 38% (n=18) reported being less likely to use a condom with a paying partner after using alcohol	
Chiao et al. 2006	Southern Philippines	19% had consumed alcohol before commercial sex 37% reported engaging in sex with inebriated customers	Alcohol consumption with a customer was significantly associated with condom use (p<0.01)	
Morisky et al. 2010	Southern Philippines	Education about HIV/AIDS decrease the daily amount of alcohol consumed	Education about HIV/AIDS increases the likelihood that a condom would be used	
Wechsberg et al. 2005	Pretoria, South Africa	18% of women reported daily alcohol use during the previous 30 days	Sexual abuse was associated with a lower chance of using a condom (p<0.01)	
Wechsberg et al. 2009	Pretoria, South Africa	FSWs were more likely that female non-SWs to be diagnosed with an alcohol or drug abuse disorder	Not discussed	
Fisher et al. 2008	Moshi, Tanzania	73.9% women in the cohort had consumed alcoholic beverages and the prevalence of problem drinking at baseline was 34.6%	Interestingly, non-drinkers were more likely to have not used a condom during their last sex	19.0% were HIV-positive HIV prevalence was 22.4% among alcohol drinkers, and 9.5% among abstainers
Fisher et al. 2010	Moshi, Tanzania	The greatest risk of condom failure (19.4%) occurred when the woman alone had been drinking (OR 14.05, CI 4.03-50.41)	The use of a condom appears to be more a function of situational negotiation and the woman's control over the outcome	
Nemoto et al. 2004	San Francisco, California, United States	67% report some alcohol consumption during the past 30 days 14% report consuming alcohol with a customer	Condoms were used consistently 91% of times for vaginal sex. Only 17% of respondents report always using a condom with their main partner	No woman reported being HIV-positive
Nemoto et al. 2008	Chi Minh City, Vietnam	89% of participants report alcohol consumption in the past year A total of 71% of FSWs report having had sex under the influence of alcohol A total of 90% of FSWs report consuming alcohol with customers	Inconsistent condom use was reported by 85% of bar/club FSWs, 72% of massage parlour FSWs, and 68% of street FSWs	Prevalence among those who had an HIV test was 7%; 18% of the street FSWs tested positive; 7% of bar/club FSWs tested positive

Table 1a. Female Sex Workers

4.2 Clients of Commercial Sex Workers (Table 1b)

In the 6 selected studies in this group many reported high levels of alcohol consumption. In one study by Havanon et al. (1993) in Thailand, 82% report that drinking accompanies commercial sex. Samet et al. (2010) for a study in Mumbai, India, shows heavy alcohol consumption was significantly associated with inconsistent condom use in male clients of FSWs (OR 2.40, CI 1.21-4.77, p=0.01). Not all selected studies in this group reported HIV status.

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Madhivanan et al. 2005	Mumbai, India	Men who report sex with a FSW, they were more likely to have consumed alcohol before sex (OR 1.5, CI 1.2-1.9, p<0.001)	Men who report sex with a FSW were more likely to have had unprotected sex (OR 1.9, CI 1.4-2.6, p<0.001)	1683 subjects had an HIV test with HIV prevalence of 14%	Heterosexual males
Samet et al. 2010	Mumbai, India	44% of men were heavy alcohol users (>14 drinks per week or >4 drinks on a single sitting)	Heavy alcohol consumption was significantly associated with inconsistent condom use in male clients of FSWs (OR 2.40, CI 1.21-4.77, p=0.01)	HIV-positive	Heterosexual males
Thomas RM et al. 1990	Edinburgh, Scotland	Over 75% (sometimes) and 40% (regularly) of subjects consumed alcohol before contacting a CSW	Alcohol use was inversely associated with condom use with male CSWs (p=0.027), but not with FSWs		Mostly male subjects: 175 had contacted FSWs; 26 had contacted MSWs; 5 had contacted both FSWs and MSWs
Dissabandara et al. 2009	Sri Lanka	54.5% used alcohol on a regular basis (weekly or more) 16.8% reported daily alcohol consumption	Of 67.3% of subjects who report having sex with a CSW, only 14.4% reported regular condom use		Heterosexual males
Havanon et al. 1993	Thailand	82% report that drinking accompanies commercial sex, while 74% of them report to have been drunk	Nearly 50% of subjects usually use condoms, and students usually use condoms 77% of times		Males
Fritz et al. 2002	Harare, Zimbabwe	84% of subjects identified themselves as alcohol drinkers	Sex while intoxicated was associated with 20 times more unprotected sex with casual partners and 27 times more unprotected sex for those paying for sex	Overall, 96 subjects were HIV-positive	Male

Table 1b. Clients of Commercial Sex Workers

4.3 Army personnel (Table 2a)

People who live for prolonged periods far from their home may be considered migrants. Among army personnel and migrant workers alcohol consumption is very high, leading to multiple sex partners, including CSWs. HIV status is provided in three of the four studies in this group. Studies on army personnel show inconsistent condom use coupled with alcohol consumption, with many in non-monogamous relationships (Brodine et al., 2003; MacQueen et al., 1996; Tavarez et al., 2010).

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Tavarez et al. 2011	Border-crossing zones on the western border of the Dominican Republic	Alcohol abuse is a predictor of sex with a CSW (OR 4.80, CI 2.00-11.30, p<0.001)	Alcohol abuse is an indicator of non-monogamous, unprotected sex (OR 2.8, CI 1.7 -4.4)		Male
MacQueen et al. 1996	Phitsanuloke Province, Northern Thailand	Alcohol consumption is described as very high in this qualitative study	Alcohol virtually eliminates the willingness of the male to use a condom	7% of men were HIV-positive	Heterosexual Males
Brodine et al. 2003	United States	64% of the cohort reported moderate to heavy alcohol consumption, defined as five drinks per session, two to three times monthly	Overall, condom use was very limited, and this can be attributed to alcohol use during sex	HIV-positive	Heterosexual Males
Conigliaro et al. 2003	Pittsburg, Pennsylvania	351 (40.40%) are present drinkers 289 (33.1%) are binge drinkers (≥6 drinks on 1 occasion) 310 (35.5%) were hazardous drinkers (binge drinkers or scored ≥8 on AUDIT)		HIV-positive	99% of the sample were male MSM made up 36% of the sample

Table 2a. Army Personnel

4.4 Migrant workers (Table 2b)

There were five studies selected for this group. One study by Gupta (2010) showed that HIV prevalence was higher among migrants than it was among non-migrants (0.60% versus. 0.33%), particularly those who drank almost every day (1.36%). Xiao et al. (2010) argue that migrants with a low level of education do not use condoms whether they bring into play alcohol or not. Verma et al. (2010) illustrated greater alcohol consumption and higher rates of unprotected sex for migrants who were away from home for prolonged periods. This study, conducted in India, reported alcohol use prior to sex in general as significantly higher in highly mobile, male migrant workers (OR 1.5, CI 1.2-1.7).

4.5 Heterosexual couples

4.5.1 In the United States with unknown HIV status (Table 3a)

Ten studies were selected for inclusion in this group. In a study by Cavanaugh et al., 2010, a woman's alcohol problem was associated with unprotected sex with a non-monogamous primary partner. In an earlier study by Lauby et al., 2001 binge drinking correlates with low condom use with both main partner and casual partners (p<0.001) Condom use was lower overall with main partner than it was with casual partners.

4.5.2 HIV-positive heterosexual couples in the United States (Table 3b)

Five studies were selected for this group. One by Theall et al. 2007 found that alcohol consumption was associated with the partner refusing to use a condom. Another, based on findings by The NIMH: Multi-site HIV/ STD-Prevention Trial for African American-Couples Group, 2010 showed that females were less likely to be alcohol dependant (9.09%) than males (14.96%), (OR 1.65, CI 1.15-2.36). Alcohol use correlated with a higher number of sex partners and lower condom use (Gerbi et al., 2009).

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Lin et al. 2005	Beijing and Nanjing, China	34.6% of the participants had been intoxicated at least once during the previous month	No association between alcohol intoxication and condom use was found		Both males and females
Gupta et al. 2010	India	6408 (8.6%) of men drink once a week 2358 (3.2%) of men drink almost daily	Condom use with a paid partner was highest for those who used alcohol almost daily (p<0.001)	HIV prevalence was higher among migrants than it was among non-migrants (0.60% vs. 0.33%), particularly those who drink almost every day (1.36%)	Both males and females
Verma et al. 2010	India	Alcohol use prior to sex in general is significantly higher in highly mobile male migrant workers (OR 1.5, CI 1.2-1.7)	Approximately one third of men report inconsistent condom use with paid partners		Heterosexual males
Amirkhanian et al. 2010	St. Petersburg, Russia	Participants report consuming 4.3 drinks per week	Mean percentage of condom use was 35.0%		Males
Rhodes et al. 2010	North Carolina, United States	Nearly 10% of the sample reported drinking alcohol every day or nearly every day Nearly 58% reported binge drinking during the past year	Alcohol consumption does not decrease condom use		Heterosexual males

Table 2b. Migrant Workers

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Lauby et al. 2001	Philadelphia and Pittsburgh (Pennsylvania), Portland (Oregon), Oakland and San Francisco (California)	Binge drinking correlates with low condom use with both main partner and casual partners (p<0.001)	Condom use was lower overall with main partner than it was with casual partners		Heterosexual females
Rosengard et al. 2005	Rhode Island Department of Corrections Women's Division	17.2% report weekly binge drinking	From 40% of women who report casual partners in the past 3 months, 47% of them report consistent condom use		Heterosexual females
O'Leary et al. 2006	A representative sample of the United States population (census data), with particular focus on Georgia, Louisiana, Alabama, Florida, South Carolina, North Carolina and Mississippi	There are significantly fewer lifelong drinkers in this region, particularly in the disproportionally-affected populations, such as African-American individuals, and young African-American females	Not discussed		Heterosexual males and females
Ryan et al. 2009	Los Angeles county, California	15 of 56 sexual events involved alcohol consumption	Condoms were used in 19 of 56 sexual events		Heterosexual females
Cavanaugh et al., 2010	New Haven, Connecticut	89.7% of women used alcohol in their life 50.7% of women have used alcohol to intoxication during the past 6 months	A woman's alcohol problem was however associated with sexual risk behaviour (OR 1.24, CI 0.46-3.54); 11.0% of women report unprotected sex with a non-monogamous primary partner		Heterosexual females
Dillon et al. 2010	Miami, Florida	Alcohol intoxication before sex was associated with younger age of oral sexual debut (p<0.01), as well as more sex partners (p<0.01)	Not discussed		Heterosexual females

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Jenness et al. 2011	New York City, New York	41.7% were binge drinkers	Overall, 38% of women had unprotected anal intercourse 47.3% of those who had unprotected anal sex in the past year were binge drinkers	9% were HIV +	Heterosexual females
Stockman et al. 2010	Nationally-representative sample of the United States	Approximately 30% of women were given alcohol or drugs prior to sexual coercion	Among 1085 women with more than 1 male sex partner in the past year, 69.0% reported no condom use at last vaginal sex		Heterosexual females
Towe et al. 2010	Baltimore, Maryland	Binge drinking during the past 30 days is linked with unprotected sex (OR1.18, CI 1.04-1.33, p=0.008)	84% reported having unprotected sex during the past 12 months Unprotected sex was higher with a main partner (84%), compared to a casual partner (63%)	10 (3%) were HIV +	Heterosexual males and female
Wilson et al. 2010	Dallas, Texas	92.2% of the sample consume alcohol; 14.4% of those reporting alcohol consumption do so more than 5 times a week	Of the 108 men who reported having had sex in the past year, 54% reported very limited or non-existent condom use. Only 17% report always using a condom		Heterosexual males

Table 3a. Heterosexual Couples in the United States with Unknown HIV Status

Study Authors	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Subject-Gender and Sexual Orientation
Stein M. et al. 2005	Brown University, Providence Rhode Island	Hazardous drinking has a statistically higher association for having any sex compared to binge drinking (p=0.0001 vs. p=0.001)	A higher probability of unprotected sex was associated with any use of alcohol	HIV-positive	67% of patients were self-identified as heterosexual
Theall et al. 2007	New Orleans, Louisiana	Approximately 25% of the women were classified as bingers	Alcohol consumption was associated with the partner refusing to use a condom (OR 1.58, CI 1.30-8.41)	HIV+; 16% report that their last partner was also HIV +	Heterosexual females; 23% report more than one male partner in the last year
Gerbi et al. 2009	Montgo-mery, Alabama	Men were significantly more likely than women to drink alcohol before sex	Alcohol use before sex was correlated with lower condom use (74% of those who do not drink before sex vs. 43% of those who drink before sex report using condoms most of the time, p=0.0001)	AIDS-positive	Heterosexual males and females
Golin et al. 2009	Seven HIV clinics in six US cities	6% of the sample report always using alcohol before sex, and 63% never using alcohol before sex	12.3% of the sample report unprotected sex with an at-risk partners in the past 3 months	HIV-positive	Heterosexual males
The NIMH: Multi-site HIV/ STD-Prevention Trial for African American-Couples Group, 2010	Atlanta, Los Angeles, New York Philadelphia	Females were less likely to be alcohol dependant (9.09%) than males (14.96%) (OR 1.65, CI 1.15-2.36)	Condom protected sex was significantly lower in females (p=0.0018)	One partner in the couple was HIV +	Heterosexual males and females

Table 3b. HIV-positive Heterosexual Couples in the United States

4.5.3 Outside of the United States (Table 3c)

Eighteen studies were included for this group: An Indian study by Kumar et al, 2010, found that current consumption of alcohol is associated with premarital sex among males only (OR 3.5, CI 2.53-4.83, $p<0.001$).l The study also found that condom use in rural areas was lower than in urban areas however this was significant only for males. In SA, violence, often of a sexual nature, perpetrated by men under the influence of alcohol, against women is of particular concern (Seedat et al., 2009). In a Cape Town study condom use was found to be low in these episodes of sexual violence (Townsend et al., 2011).

Study	**Study Settings**	**Alcohol Use Pattern**	**Condom Use Pattern**	**HIV Status**	**Gender and Sexual Orientation of Subjects**
Dandona et al. 2008	Guntur district, Andhra Pradesh, India	Consuming alcohol before sex was associated with HIV in males only (OR 3.60, CI 2.08-8.22)	Not discussed	Not quantified	Heterosexual males and females
Thomas BE et al. 2009	Chennai, India	16.7% of men and 0% of women report alcohol consumption	More than 80% of men 100% of women reported that they did not use condoms	HIV-positive	Heterosexual males and females
Berg et al. 2010	Navi Mumbai, India	Drinking patterns ranged from social drinking to overindulgent and heavy drinking	Not discussed		Heterosexual males
Kumar et al. 2011	Guntur district, Andhra Pradesh, India	Current consumption of alcohol is associated with premarital sex among males only (OR 3.5, CI 2.53-4.83, $p<0.001$)	Condom use in rural areas was lower than in urban areas, however this was significant only for males		Heterosexual males and females
Nayak et al. 2010	Karnataka, India	Among drinkers, the mean quantity of alcohol consumed was 60 g (5 drinks) 15% of the sample used alcohol 4 or more times a week	Not discussed		Heterosexual males
Singh SK et al. 2010	Navi Mumbai, India	Mean thirty day mL of alcohol consumed was approximately 230 and less than 5% of drinkers, drank over 1000 mL a month	High alcohol consumption correlates with high condom use ($p<0.002$)		Heterosexual males
Hargreaves et al, 2002	Kisumu, Kenya	For both men (48%) and women (15%), highest alcohol consumption in the past month was among those 25-49 years with a higher socio-economic status	Condom use was higher among males (31.1%) and females (16.8%) 15-24 years, with a higher socio-economic status	HIV prevalence was 19.8% in males and 30.2% in females	Heterosexual males and females
Raj A et al. 2009(b)	St. Petersburg, Russia	72% reported alcohol consumption during the past 30 days 64% reported binge drinking – 93% of these were alcohol dependent	88% report unprotected sex with a main partner, while 76% report unprotected sex with a casual partner	8% of binge drinkers and 28% of non-binge drinkers were HIV+ ($p<0.001$)	Heterosexual males and females
Kalichman et al. 2006	Cape Town, South Africa	Alcohol consumption in the context of sex was higher for individuals reporting a drinking problem (OR 24.4, CI 14.3-41.4, $p<0.01$) and for individuals whose partner reported a drinking problem (OR 5.1, CI 3.1-8.5, $p<0.01$)	No association between problem drinking and condom use was found		Heterosexual males and females

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Wechsberg et al. 2008	Cape Town, South Africa	At baseline, Black women report alcohol use on more days in the past month than Colored women (13.55 vs. 5.82, p<0.001)	Black women were more likely to use a condom (50% vs. 15%) and to have only one partner (main)		Heterosexual females
Wong et al. 2008	Cape Town, South Africa	High levels of problem drinking were found among both men (58%) and women (42%)	Not discussed		Heterosexual males and females
Andersson et al. 2009	Soweto, South Africa	Males reported higher rates of heavy alcohol use (p<0.001)	Substance use and male gender predicted higher risk behaviours, including inconsistent condom use	HIV-negative	Heterosexual males and females
Avalos et al. 2010	Cape Town, South Africa	Hazardous alcohol use was reported by 12.6% of the entire sample	Multiple sex partners and inconsistent condom use were the two most common sexual risk behaviours reported		Heterosexual males and females
Kalichman et al. 2010	Cape Town, South Africa	In a multivariate analysis, alcohol use was significantly associated with unprotected sex with serodiscordant partners (OR 2.5, CI 1.0-6.5, p<0.05)	Condom use was generally high for HIV-positive individuals, however it was low when engaging in seroconcordant sex	1479 (85%) of the sample was tested for HIV 218 (12%) of those tested were HIV +	Heterosexual males and females
Townsend et al. 2011	Cape Town, South Africa	Problem alcohol use was associated with both physical and sexual intimate partner violence, with inconsistent condom use	Physical intimate partner violence was associated with inconsistent condom use		Heterosexual males
Kapiga et al. 1998	Dar es Salaam, Tanzania	29.5% of subjects consumed alcohol	Condom use was not different between drinkers and non-drinkers	All women started off HIV-negative Drinking was associated with risk of HIV (OR 2.43, CI 1.54-3.82)	Heterosexual females
Ghebremichael and Paintsil, 2009	Moshi district, Tanzania	About 33% of the participants were categorized as alcohol abusers	80% of subjects never used condoms in the past 12 months, and only 6% used them often	6.5% of men were HIV-positive	Heterosexual males
Emusu et al. 2009	Kampala, Jinja and Mbale, Uganda	Alcohol abuse by HIV-positive male partners often resulted in them perpetrating unprotected sex	Unprotected	About half of the women were HIV-negative (and their male partners were HIV +), and the other half were HIV+ (and their male partners were HIV-negative)	Heterosexual females

Table 3c. Heterosexual Couples Outside of the United States

4.6 Men Who Have Sex with Men (MSM) (Table 4)

Drug and alcohol abuse is reported to be particularly high among men who have sex with men; this has implications for risky sexual practices. Eleven studies were selected for inclusion in this group, some studies included only HIV-positive and some HIV-negative participants. In one study by Koblin et al., 2006, of initially HIV-negative persons, 29% of sero-conversions in the cohort was attributed to alcohol use (odds ratio 2.54, CI 1.83 to 3.53). In one study, MSM were identified as engaging in unprotected sex almost three times as often as any other group, including IDUs (Stein et al., 2005). Like the American MSM, the French MSM are also more likely to engage in unprotected sero-discordant sex following binge alcohol consumption (Bouhnik et al., 2007). While protected sex is higher with casual partners than with regular partners for MSM, alcohol consumption further lowers the likelihood of protected sex with a casual partner Unprotected sex with casual partners is associated with a greater risk of frequent alcohol use before or during sex (OR 1.5, CI 1.03-2.24, p=0.037), (Folch et al., (2009))

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Prestage et al. 2009	Australia	Drugs and alcohol are used to enhance the sexual experience in more 'adventurous' gay community subcultures	Having more than five drinks (OR 2.41, CI 1.34-4.33, p=0.003) was associated with unprotected anal intercourse with non-HIV sero-concordant partners		MSM
Lambert et al. 2009	Montreal, Canada	Alcohol was used before sex 39.4% of the times when the partner was regular and 49.8% of the times when the partner was casual	Alcohol use was associated with unprotected anal sex at last sexual episode 12.2% of participants had unprotected anal sex at last sexual episode	Self-reported HIV-negative or HIV status-unknown	MSM
Tripathi et al. 2009	Tallinn, Estonia	Mean alcohol consumption was 7.1 standard alcoholic drinks per week in the week preceding the study	Over 50% of the sample did not use a condom regularly in the past 12 months Higher alcohol consumption was negatively associated with use of condom during the last intercourse		MSM
Bouhnik et al. 2007	France	Alcohol consumption at least once a month, is associated with unprotected sero-discordant sex (OR 2.4, CI 1.2-4.9, p=0.003)	Unprotected sex was more prevalent within sero-concordant couples than it was in sero-discordant couples	AIDS-positive	MSM
Tsui and Lau, 2010	Hong Kong	Only 13.3% of subjects reported that they drank alcohol before sex	MSM who chose their sexual partners from the internet are more likely to engage in unprotected sex		MSM
Mendoza-Pérez et al. 2009	Ciudad Juárez, Chihuahua, Mexico	29.6% report consuming alcohol more than twice a week	Alcohol consumption was associated with engaging in unprotected sex		MSM
Lane et al. 2009	Soweto, South Africa	87.9% report that they drank at least once per month 75.9% scored positive for problem drinking 54.5% of subjects report 10 or more drinks on a typical day of drinking	Unprotected anal intercourse predicts HIV-positive status (28.0%, CI 21.9%-33.6%)	HIV prevalence was estimated at 13.2% Problem drinking predicts HIV-positive status (75.9%, CI 70.0%-82.1%)	MSM

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Folch et al. 2009	Catalonia, Spain	19.6% of men were frequent users of alcohol 63.8% of men used alcohol and drugs at some time before or during sex	Unprotected sex with casual partners is associated with a greater risk of frequent alcohol use before or during sex (OR 1.5, CI 1.03-2.24, p=0.037)	24.3% of frequent alcohol users were HIV-positive	MSM
van Griensven et al. 2010	Bangkok, Thailand	80.2% of MSM report alcohol consumption in the past 4 months Of these, 11.5% binge drank	Always using a condom was reported by 43.7% of MSM, and this increased to 70.8% when the partner was paid	HIV-negative	MSM
Stein M. et al. 2005	Providence, Rhode Island	Alcohol use (OR 1.95, p<0.01) and hazardous alcohol use (OR 2.90, p<0.01) were associated with having sex	Alcohol use (OR 2.30, p<0.01) and hazardous alcohol use (OR 2.66, p<0.01) were associated with unsafe sex	HIV-positive	Heterosexual males and females and MSM
Koblin et al. 2006	Boston (Massachusetts), Chicago (Illinois), Denver (Colorado), New York (New York), Seattle (Washington), San Francisco (California)	72.1% of men reported using alcohol or drugs before having sex	69.1% report any unprotected anal intercourse	Initially HIV-negative, 29% of seroconversions in the cohort was attributed to alcohol use (odds ratio 2.54, CI 1.83 to 3.53)	MSM
Golin et al. 2009	Denver (Colorado), Kansas City (Missouri), Nashville (Tennessee), Brooklyn (New York), Chapel Hill (North Carolina) and Atlanta (Georgia)	26% of MSM report binge drinking at least once a week 51% of MSM report drugs or alcohol use before sex in the past 3 months	23.0% of MSM report unprotected sex	HIV-positive	MSM
Operario et al. 2009	Oakland, California	70.6% of subjects report alcohol use before sex in the past 30 days	33.8% to 51.5% had unprotected anal sex with a male Alcohol use was not linked to unprotected sex	17.6% reported being HIV-positive	MSM who report sexual relationships with females
Mackesy-Amiti et al. 2010	Chicago, Illinois	93% individuals used alcohol 39% individuals showed signs of alcohol dependence	Between 28% and 35% of the sample report unprotected anal intercourse	HIV-negative	MSM
Reisner et al. 2010	Massachusetts	Overall, 29% of the sample was found to abuse alcohol at the time of enrolment into the study	Problem drinking was linked to unprotected anal sex with a sero-discordant male partner (OR 3.22, CI 1.22-8.50, p<0.02)		MSM
Semple et al. 2010	San Diego, California	The average number of drinks in a typical drinking day was 3.5	The average number of unprotected anal sex acts in the past 2 months was 10	HIV-positive	MSM

Table 4. Men Who Have Sex with Men

4.7 Injecting Drug Users (IDUs) (Table 5)

For injecting drug users, alcohol consumption leads to needle sharing, unsafe sex and exchanging sex for money and drugs, opening the door for HIV transmission and re-infection. Nine studies were selected for this group with mixed HIV status. A study in

Nepal conducted by Poudel et al. 2010 reported that 44% of non-drinkers share needles, as opposed to 55% of alcohol drinkers (OR 0.63, CI 0.38-1.03). For IDUs, condom use was higher when both partners were sober and it was higher overall when the partner was a casual partner, as opposed to a main partner (Arasteh and Des Jarlais, 2009). There also seems to be an association between the severity of substance abuse and the degree of risk taking resulting in contracting STDs (Chan et al., 2010).

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern	HIV Status	Gender and Sexual Orientation of Subjects
Poudel et al. 2010	Kathmandu Valley, Nepal	44% of non-drinkers share needles, as opposed to 55% of alcohol drinkers (OR 0.63, CI 0.38-1.03)	Not discussed	21% of 202 participants who had taken an HIV test tested positive	Males
Matos et al. 2004	Vega Baja, greater San Juan, Puerto Rico	Alcohol intoxication during the last 30 days was reported by 18% of participants	Those reporting alcohol intoxication were also more likely to report unprotected sex with a paying partner and with a casual partner		Both males and females
Parry et al. 2008(a)	Cape Town, Durban and Pretoria, South Africa	Alcohol use is not mentioned	Being high was a reason to not think about safe (have unprotected sex)	28% of individuals who offered to be tested for HIV tested positive	Mostly heterosexual male and female
Parry et al. 2008(b)	Cape Town, Durban and Pretoria, South Africa	Alcohol is not mentioned	Drugs led to inconsistent condom use	Among MSM who agreed to HIV testing, one-third tested positive	Both heterosexual males and MSM
Stein MD. et al. 2000	Providence, Rhode Island	60% had used alcohol in the last month 14% were at-risk alcohol abusers	Not discussed	89% of subjects had been tested for HIV, and 4% of these subjects tested HIV-positive	Males and females
Arasteh and Des Jarlais, 2009	New York City, New York	35% of the sample were at-risk drinkers (defined as more than 14 drinks per week for males or 7 drinks per week for females)	At-risk drinkers report lower condom use than occasional or non-drinkers	HIV-positive	Both males and females
Chan et al. 2010	67 cities in 29 states across the United States	35.2% of males and 44.8% of females (37.9% average) had sex while high on alcohol or drugs ($p<0.001$)	33.3% of males and 44.3% of females (36.4% average) had sex without protection ($p<0.001$)		Heterosexual males and females
Jenness et al. 2010	New York City, New York	34.5% of participants binged on alcohol at least weekly in the past year	IDU partnerships were associated with risky unprotected sex	Overall, 7.0% tested positive for HIV	Heterosexual males and females
Sander et al, 2010	Baltimore, Maryland	At study entry, 36% of participants were binge drinkers	Not discussed	HIV-negative at the start of the study	Predominantly heterosexual males and females (MSM was reported at 1% of visits)

Table 5. Injecting Drug Users

4.8 Emerging adults

In emerging adult populations, alcohol is often consumed in the context of parties, with the potential for multiple sex partners, and unsafe sexual encounters.

In United States (Table 6a) Ten studies showed that outcome variables of interest differed across race, age and sex. (Hou, 2009) established that Black students were less likely to use alcohol before any type of sexual activity. However this group was more likely to use a condom during sex of any kind. Adefuye et al, (2009) found that condom use was higher in younger individuals when compared with older individuals. Alleyne et al., (2010) noticed that alcohol use at last sex was 18.0% overall, with 15.0% in females and 21.1% in males. Also males were more likely to have used a condom during their last sexual episode.

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern
Sheer and Cline, 1995	A large South-Eastern University, United States	Alcohol is often consumed at parties	No significant association was found between the number of sex partners and condom use
Apostolopoulos et al. 2003	Sample was representative of university students, United States	Alcohol or drugs influenced decisions involving sex	31% of males and 32% of females report that alcohol use prior to sex had a negative influence on their decision to use a condom
Morrison et al. 2003	Seattle, Washington	Most subjects had been drinking more than one or two drinks per occasion and most have been getting drunk at least several times per week	The odds of condom use were not associated with whether or not alcohol was consumed before sex. Condom use was marginally lower for females than it was for males (OR 0.30, CI 0.11-1.06)
Adefuye et al. 2009	An urban Midwestern university, United States	14.6% of individuals below the age of 19 report alcohol use before sex	Condom use was higher in younger individuals
Hou, 2009	Southern universities, United States	Black students were less likely to use alcohol before any type of sexual activity	Black students were more likely to use a condom during sex of any kind
Murphy et al. 2009	Data collected from the National Longitudinal Survey of Youth	Males were found to consume higher levels of alcohol	The highest alcohol consumption was associated with the group that showed the highest sexual risk tendencies in both males and females
Randolph et al. 2009	A Southern University, United States	62.9% of participants reported binge drinking	Condom use was greater among older participants. Higher perceived risk of HIV was also associated with greater condom use
Alleyne et al. 2010	Youth Risk Behavior Survey, Illinois	Alcohol use at last sex was 18.0% overall, with 15.0% in females and 21.1% in males	Males were more likely to have used a condom during their last sexual episode.
Nkansah-Amankra et al. 2010	Colorado Youth Risk Behavioral Survey, Colorado	Alcohol consumption was found to lead to sexual risk taking and multiple sex partners	Not discussed
Wells et al. 2010	New York, New York	62.9% reported to sex after use of alcohol	Not discussed

Table 6a. Emerging Adults in the United States

Outside of the United States (Table 6b) HIV status is provided in two of the seven selected studies. Each of these seven studies provides valuable insight into the dynamics operating within this high-risk group relating to alcohol consumption and unsafe sex. A Chinese study (Xiao et al., 2010), showed that using alcohol diminished the likelihood of participants using condoms. In a South African study by McGrath et al., 2009, it was observed that for men, age at first sex was associated with having ever used alcohol (OR 1.89, CI 1.55-2.30, p<0.001).

Study	Study Settings	Alcohol Use Pattern	Condom Use Pattern
Tavares et al. 2009	Santiago Island, Cape Verde	Alcohol consumption was associated with sexual initiation in males	84.72% reported having used a condom in their first sexual intercourse
Xiao et al. 2010	Hunan Province, China	Using alcohol diminished the likelihood of participants using condoms	Impulsivity was shown to be negatively correlated with condom use
Campo-Arias et al. 2010	Santa Marta, Colombia	Of 804 students with a history of sexual relationships, 18.4% reported sexual intercourse after alcohol consumption	33.7% informed inconsistent condom use
Karnell et al. 2006	Pietermaritzburg area, KwaZulu-Natal, South Africa	At pre-test, almost half of the sample (both males and females) were using alcohol	A greater willingness to use a condom during the next 3 months was reported by students in the intervention group (p<0.01)
Morojele et al. 2006	Grade 8 and grade 11 students in Cape Town, South Africa	Some boys report that they supplied alcohol to girls in order to sleep with them. Some girls report that they consumed alcohol and other drugs voluntarily, and this made them more prone to engage in sexual acts	Some boys avoided condom use in order not to decrease sexual pleasure
McGrath et al., 2009	KwaZulu-Natal, South Africa	In men, age at first sex was associated with having ever used alcohol (OR 1.89, CI 1.55-2.30, p<0.001). The same trend was seen in women, although it was not statistically significant	Not discussed
Singh K et al., 2010	Hwange District, Zimbabwe	61% of individuals aged 15 to 19 that were found in nightlife/drinking venues report ever to have had sex – this was significantly higher than individuals in any other type of venue (p<0.01)	Condom use at last sex was 53.9% in the sample reporting the highest number of sexual partners

Table 6b. Emerging Adults Outside of the United States

4.9 Condom use

Not all studies have consistent findings with regard to condom use and alcohol consumption. The first important criterion may have to do with how they are used, rather than merely whether condoms are used. The fact that condom use is documented does not automatically imply that the condom was used effectively, i.e. there were no accidents including breakage and spillage, and that a condom was used for every separate act of sexual intercourse that took place during a particular sexual encounter.

Effective condom application for every act of sexual intercourse is a key means of limiting the spread of HIV. Condoms not only prevent HIV infection and other STIs but also prevent unplanned pregnancies. Condom use is recommended for monogamous couples if sero-discordant or both are infected. The number and sequencing of the sexual acts as well as the characteristics of the persons and nature of the sexual behaviour involved in the partnerships are not that relevant if there is adequate protection for contracting STIs.

Information concerning condom use for protection against contracting HIV is known but not necessarily acted upon. Hence an important focus of public health efforts should be on addressing other intermediary risk factors for sexual HIV risk behaviour, particularly those risk factors resulting from the use of alcohol.

Basic research on condom use and number of sexual partners is stymied by social desirability bias. The data obtained on safe-sex practices may be valid or there may be over-estimates on condom usage as well as under or over-estimates on the number of sexual partners. Furthermore, it is necessary to refer to the mechanics of condom use when inebriated. A South African study by Townsend et al. (2010) attempted to answer the question: ''Are condoms applied less effectively and consistently by men who drink heavily compared to those who do not?'' The study replicated findings from other studies, (Kalichman et al, 2008; Simbayi et al. 2006) that heavy drinkers, if they use condoms, use them inconsistently.

4.10 Concurrency

Johnson et al. (2009) applied a mathematical model that demonstrated that concurrency is a major driver of the HIV/AIDS epidemic in SA. Concurrency in sexual relationships refers to sexual encounters that overlap in time with different partners, usually two or more simultaneous relationships. The role of concurrency in the spread of HIV is not straight forward. Concurrency subsumes various other factors such as the levels of infectiousness of HIV, so it may be particularly relevant to the spread of HIV, for instance early on in the epidemic or for a newly infected person. According to Morris 2010 concurrency increases the risk of transmitting HIV not acquiring it. One may debate whether long term polygamous relationships or several monogamous relationships in quick succession have a greater impact on HIV transmission. Morris argues that as the connectivity in sexual networks is non–linear a small behavioural change can result in significant HIV prevention. She suggests that if 5% of those individuals who are sexually active have their partners serially rather than concurrently, without reducing the number of partners, this will impact positively on the HIV epidemic.

The Townsend (2010) study referred to above showed high rates of problem drinking for men who have multiple, concurrent sexual partners (It should be noted that the figure from Townsend (2010) for problem-drinkers of 58.5% for men in urban contexts in the Western Cape (South Africa) is considerably higher than the 27.9 % of lifetime problem drinkers in the same region recorded in the first South African Demographic and Health Survey (Parry et al. 2005). Despite this finding, Townsend (2010) did not find that problem drinkers were more likely to have multiple partners than non-problem drinkers. Kalichman et al. (2007) however showed that greater frequency and quantities of alcohol use was related to a greater number of sexual partners. Furthermore, an interesting finding from the Townsend (2010) study is that the amount of alcohol consumed effects the choice of sexual partner, i.e. one who is more likely to drink alcohol and be unemployed; this in turn has additional implications for riskier sexual encounters. For example, with transactional sex there is less likelihood of condom use.

The link between alcohol use, condom use and concurrency is not clear-cut; no pattern is discernable both within and between groups. A somewhat counter intuitive finding is that condoms are more likely to be used for casual partners, i.e. conferring protection against HIV infection.

4.11 Limitations of the study

Overall, this study has limitations in that only English language papers were considered and comprised mainly of those published between 2008 and 2010. Furthermore, HIV status of the subjects was not always provided in the papers and alcohol use patterns were often loosely defined. The emerging adult high risk group may also include some studies that can be considered to constitute a separate high risk group, namely adolescents, for example, Morojele 2006b. Groups are not mutually exclusive, individuals within groups may also be part of other groups, for example men who classify themselves as heterosexual may also engage in sex with other men.

Katz (2008) highlights that there are significant risk factors that are not captured from standard behavioural indicators, Katz also suggests that cumulative risk, that is, risk over an extended period is not captured by annual measures of sexual risk, such as the number of sexual partners in the past 12 months. These concerns impact on the adequacy of the measures of sexual risk behaviour and may explain why the demonstrated behavioural changes (Shisana et al., 2009) have not lead to the expected declines in HIV measures in SA. For example, despite evidence showing increased condom use there is an indication that condoms are not used consistently.

4.12 Models

Epidemiological models combine epidemiological and statistical data, such as the probability of contracting HIV per sexual encounter, as well as behavioural data to help elucidate the spread and the control of HIV. Models are necessary to simulate complex sexual networks. For example, Wim Delva's SIMPACT is a modeling tool used to simulate HIV epidemics in complex sexual networks. SIMPACT can model the effect of concurrency on the pathogenesis of HIV/AID. For example by "capturing the formation and dissolution of sexual relationships between individuals" it can demonstrate the impact of this interpersonal behaviour on HIV transmission. (Delva thesis. Page 120). Wim Delva argues that "the potential for epidemiological models to improve our understanding of the determinants of HIV spread and control may only be fully unlocked when questions about the sexual network structure and partnership dynamics are adequately addressed both by empirical studies and ensuing advances in model development." (Delva thesis. Page 120)

"SIMPACT is a rather flexible tool, so in principle it is indeed possible to model the effect of alcohol (reduction) on the transmission of HIV in the population. However, the effect of alcohol works indirectly through behavioural choices (primarily formation of sexual relationships and condom use). If data are available describing how alcohol impacts on sexual risk behaviour, SIMPACT will be able to simulate what the implications of these effects are in terms of enhanced HIV transmission." (e-mail correspondence from Wim Delva)

5. Conclusion and recommendations

This literature review again demonstrated the strong association between alcohol consumption and HIV transmission via unsafe sex; broadly defined as unprotected sexual intercourse. This literature review examined two HIV prevention strategies, condom use and concurrency with the additional behavioural component, namely, alcohol consumption at time of sexual encounter in selected high risk sub-populations. It shows that the link between drinking and unsafe sex, that is, inconsistent condom use and multiple sexual partners, is influenced by diverse factors such as the amount of alcohol consumed (generally but not always), the setting and power relations, among other variables and that this holds across the risk groups studied and in different countries. To a large extent this literature review substantiates the findings of previous reviews in this area. Although it does not close the question of whether alcohol use is causally related to unsafe sex, it does add more evidence to the established association. It extends the area of research in that it focuses on high risk groups globally. In so doing it indicates important future areas of research, namely a focus on vulnerable sub-populations and the necessity to intervene on alcohol consumption and its role in leading to risky sex and subsequent sero-conversion to HIV.

In Africa and other countries where the HIV epidemic is mainly driven by unprotected sexual intercourse, focusing on factors influencing sexual risk behaviour is paramount in preventing new and re-infections of HIV. The focus on high-risk groups is not to further stigmatise certain groups that include drug users, sex workers and MSM. The focus on high-risk behaviours in these groups is vital in order to pinpoint group specific HIV prevention interventions.

Specific recommendations for additional studies in SA in the identified high-risk groups from this review are the military and migrant workers in SA. The focus should be on alcohol consumption and risky sexual practices within these high-risk groups. The studies of the SA military should include SA peacekeepers in Africa on their return to SA. There is also a need for more studies on migrants to SA from the rest of Africa. Oscillating internal rural / urban migration was the cornerstone of Apartheid labour policies. Lurie (2010) suggests that labour migration early on in the HIV epidemic in SA, i.e. in the early 1990's, was critical to the dissemination of the virus from urban to rural areas. However, in the light of evidence of HIV transmission in rural areas he questions the uni-directionality of the spread of HIV and recommends interventions aimed at migrants and their partners to limit the transmission of HIV. Another mobile population to target for HIV prevention measures are truck drivers and the CSWs at truck stops, Ramjee and Gouws, 2002 conclude that truckers may have facilitated the spread of HIV infection in southern Africa.

Using condoms, HIV counselling and testing, and needle and syringe programs were all found to be effective and cost-effective techniques to prevent HIV infection among IDUs in a review of Thai literature (Pattanaphesaj and Teerawattananon, 2010). As for IDUs in SA, Parry (2010a) examines risk behaviour in both IDUs and non-injection drug use (NIDU), vulnerable populations at risk for HIV, to inform service delivery that includes prevention and harm reduction. Although the focus of this study is substance abuse that does not include alcohol, many of the recommendations apply equally to alcohol abuse. These include a specific recommendation that service delivery be more integrated, with HIV counselling and testing be provided at substance abuse centres, and that VCT address both substance abuse and sexual risk behaviours. An important recommendation that is pertinent to the current review is that VCT be more specific to sub-populations. In South Africa, sex work is illegal and this adds to the challenge of identifying and implementing HIV prevention interventions.

It is important to identify places for HIV prevention programmes to be developed and implemented. For example, shebeens or township taverns are generally frequented by mixed social networks of mainly men who congregate to drink, socialise and find new, usually casual sexual partners (Morojele 2006a). An intervention in a bar with peer leaders has been shown to reduce risky sex for patrons in gay bars (Kelly et al. 1997). Furthermore, the type of research on reduced HIV risk needs to encompass alcohol and other substance abuse as well as poor mental health and social problems; particularly in vulnerable sub-populations. (Jewkes 2010a; Jewkes 2010b; Sikkema 2011).

Parry et al. (2010b) states that in order to address poor health outcomes resulting from alcohol-related risky sex or non-adherence to antiretroviral regimens resulting from alcohol misuse, primary health care level personnel should use appropriate instruments to screen for problem alcohol use and if necessary provide brief interventions for substance abuse and VCT for HIV. The latter includes proper training of health and social services workers. Overall there should be greater emphasis on advocacy concerning the negative health outcomes from alcohol misuse. This includes educating people on the link of alcohol to unsafe sexual practices and the re-infection/infection with, transmission and progression of HIV/AIDS.

Intervention research should include an examination of the efficacy and cost-effectiveness of Brief Interventions for problem alcohol use as HIV prevention. Epidemiological and operational research is needed specifically on the prevalence of alcohol problems in patients with HIV and the integration of alcohol and HIV services in the public health sector at primary health care level (Parry 2010b).

The present review clearly shows that there are specific groups at higher risk of contracting HIV. These groups would benefit from improved risk-assessment information. To this end we suggest additional studies in this domain utilising meta-analyses and the modelling of data obtained for the specific groups to better summarise and utilise the information obtained. The statistical data from this review can be utilised as parameter values for modeling the spread of HIV in these sub-populations. In addition to the SIMPACT model there are others, including the STDSIM model (Habbema, 1996) that utilises different data sources to determine which prevention and intervention programmes for STIs are most cost effective.

Evidence–based, cost effective HIV interventions for specific high risk groups are a priority. In order to be optimally effective these should take into account the norms and practices central to the particular sub-population. Broader structural problems, outside the public health ambit, such as gender inequality and poverty need to be addressed in the longer term, as they are the underlying, albeit more distal risks factors resulting in the spread of HIV in SA.

6. Acknowledgement

We acknowledge the U.S. President's Emergency Fund for AIDS Relief (PEPFAR) through the US Centers of Disease Control and Prevention (CDC) (PO S-SF750-06-M-0781), and the South African Medical Research Council for funding some of the activities that informed this manuscript. Its contents, however, are solely the responsibility of the authors and do not necessarily represent the official views of the CDC, PEPFAR or other organizations referred to above.

7. Abbreviations

AIDS - acquired immune deficiency syndrome
ARV - antiretroviral drugs

CSW - commercial sex workers
DALYs - disability adjusted life years
FSW - female sex workers
HIV - human immunodeficiency virus
IDU - Injecting Drug Users
MARPs- most at-risk populations
MSM – men having sex with men
SA - South Africa
SSA - sub-Saharan Africa
STDs –sexually transmitted disease
STIs – sexually transmitted infections
US – United States of America

8. References

Adefuye AS, Abiona TC, Balogun JA, and Lukobo-Durrell M. 2009. HIV sexual risk behaviors and perception of risk among college students: implications for planning interventions. *BMC. Public Health* 9: 281.

Alleyne B, Coleman-Cowger VH, Crown L, Gibbons MA, and Vines LN. 2011. The effects of dating violence, substance use and risky sexual behavior among a diverse sample of Illinois youth. *J Adolesc.* 34 (1): 11-18.

Amirkhanian YA, Kuznetsova AV, Kelly JA, Difranceisco WJ, Musatov VB, Avsukevich NA, Chaika NA, and McAuliffe TL. 2010. Male Labor Migrants in Russia: HIV Risk Behavior Levels, Contextual Factors, and Prevention Needs. *J Immigr. Minor. Health.*

Andersson KM, Van Niekerk RM, Niccolai LM, Mlungwana ON, Holdsworth IM, Bogoshi M, McIntyre JA, Gray GE, and Vardas E. 2009. Sexual risk behaviour of the first cohort undergoing screening for enrollment into Phase I/II HIV vaccine trials in South Africa. *Int. J STD AIDS* 20 (2): 95-101.

Apostolopoulos Y, Sonmez S, and Yu CH. 2002. HIV-risk behaviours of American spring break vacationers: a case of situational disinhibition? *Int. J STD AIDS* 13 (11): 733-743.

Arasteh K, and Des Jarlais DC. 2009. At-risk drinking and injection and sexual risk behaviors of HIV-positive injection drug users entering drug treatment in New York City. *AIDS Patient. Care STDS.* 23 (8): 657-661.

Avalos LA, Mertens JR, Ward CL, Flisher AJ, Bresick GF, and Weisner CM. 2010. Stress, substance use and sexual risk behaviors among primary care patients in Cape Town, South Africa. *AIDS Behav.* 14 (2): 359-370.

Baliunas D, Rehm J, Irving H, and Shuper P. 2010. Alcohol consumption and risk of incident human immunodeficiency virus infection: a meta-analysis. *Int. J Public Health.* 55 (3): 159-166.

Berg MJ, Kremelberg D, Dwivedi P, Verma S, Schensul JJ, Gupta K, Chandran D, and Singh SK. 2010. The effects of husband's alcohol consumption on married women in three low-income areas of Greater Mumbai. *AIDS Behav.* 14 Suppl 1: S126-S135.

Bouhnik AD, Preau M, Schiltz MA, Lert F, Obadia Y, and Spire B. 2007. Unprotected sex in regular partnerships among homosexual men living with HIV: a comparison between sero-nonconcordant and seroconcordant couples (ANRS-EN12-VESPA Study). *AIDS* 21 Suppl 1: S43-S48.

Bowen KM, Dzuvichu B, Rungsung R, Devine AE, Hocking J, and Kermode M. 2010. Life Circumstances of Women Entering Sex Work in Nagaland, India. *Asia Pac. J Public Health*

Braithwaite RS, Conigliaro J, Roberts MS, Shechter S, Schaefer A, McGinnis K, Rodriguez MC, Rabeneck L, Bryant K, and Justice AC. 2007. Estimating the impact of alcohol consumption on survival for HIV+ individuals. *AIDS Care* 19 (4): 459-466.

Brodine SK, Starkey MJ, Shaffer RA, Ito SI, Tasker SA, Barile AJ, Tamminga CL, Stephan KT, Aronson NE, Fraser SL, Wallace MR, Wegner SA, Mascola JR, and McCutchan FE. 2003. Diverse HIV-1 subtypes and clinical, laboratory and behavioral factors in a recently infected US military cohort. *AIDS* 17 (17): 2521-2527.

Campo-Arias A, Ceballo GA and Herazo E. 2010. Prevalence of pattern of risky behaviors for reproductive and sexual health among middle- and high-school students. *Rev. Lat. Am. Enfermagem.* 18 (2): 170-174.

Cavanaugh CE, Hansen NB, and Sullivan TP. 2010. HIV sexual risk behavior *among low-income women experiencing intimate partner violence: the role of posttraumatic stress disorder. AIDS Behav.* 14 (2): 318-327.

Chan YF, Passetti LL, Garner BR, Lloyd JJ, and Dennis ML. 2011. HIV risk behaviors: risky sexual activities and needle use among adolescents in substance abuse treatment. *AIDS Behav.* 15 (1): 114-124.

Chersich MF, Luchters SMF, Malonza IM, Mwarogo P, King'olaand N, Temmerman M. 2007. Heavy episodic drinking among Kenyan female sex workers is associated with unsafe sex, sexual violence and sexually transmitted infections. *Int J STD AIDS* 18:764-769.

Chersich MF and Rees HV. 2010. Causal links between binge drinking patterns, unsafe sex and HIV in South Africa: it's time to intervene. *Int. J STD AIDS 21* (1): 2-7.

Chiao C, Morisky DE, Rosenberg R, Ksobiech K and Malow R. 2006. The relationship between HIV/Sexually Transmitted Infection risk and alcohol use during commercial sex episodes: results from the study of female commercial sex workers in the Philippines. *Subst. Use. Misuse* 41 (10-12): 1509-1533.

Conigliaro J, Gordon AJ, McGinnis KA, Rabeneck L, and Justice AC. 2003. How harmful is hazardous alcohol use and abuse in HIV infection: do health care providers know who is at risk*? J Acquir. Immune. Defic. Syndr.* 33 (4): 521-525.

Cook RL, and Clark DB. 2005. Is there an association between alcohol consumption and sexually transmitted diseases? A systematic review. *Sex Transm Dis.* 32 (3): 156-164.

Dandona L, Dandona R, Kumar GA, Reddy GB, Ameer MA, Ahmed GM, Ramgopal SP, Akbar M, Sudha T, and Lakshmi V. 2008. Risk factors associated with HIV in a population-based study in Andhra Pradesh state of India. *Int. J Epidemiol.* 37 (6): 1274-1286.

Delva W. 2010. *Sexual behaviour and the spread of HIV. Statistical and epidemiological modelling applications.* Ghent University. Doctoral thesis.

Dillon FR, De La Rosa M, Schwartz SJ, Rojas P, Duan R, and Malow RM. 2010. US Latina age of sexual debut: long-term associations and implications for HIV and drug abuse prevention. *AIDS Care* 22 (4): 431-440.

Dissabandara LO, Dias SR, Dodd PR, and Stadlin A. 2009. Patterns of substance use in male incarcerated drug users in Sri Lanka. *Drug Alcohol Rev.* 28 (6): 600-607.

Emusu D, Ivankova N, Jolly P, Kirby R, Foushee H, Wabwire-Mangen F, Katongole D, and Ehiri J. 2009. Experience of sexual violence among women in HIV discordant unions after voluntary HIV counselling and testing: a qualitative critical incident study in Uganda. *AIDS Care* 21 (11): 1363-1370.

Fisher JC, Bang H, and Kapiga SH. 2007. The association between HIV infection and alcohol use: a systematic review and meta-analysis of African studies. *Sex Transm Dis.* 34 (11): 856-863.

Fisher JC, Cook PA, Sam NE, and Kapiga SH. 2008. Patterns of alcohol use, problem drinking, and HIV infection among high-risk African women. *Sex Transm Dis.* 35 (6): 537-544.

Fisher JC, Cook PA, and Kapiga SH. 2010. Alcohol use before sex and HIV risk: situational characteristics of protected and unprotected encounters among high-risk African women. *Sex Transm Dis.* 37 (9): 571-578.

Folch C, Esteve A, Zaragoza K, Munoz R, and Casabona J. 2010. Correlates of intensive alcohol and drug use in men who have sex with men in Catalonia, Spain. *Eur. J Public Health* 20 (2): 139-145.

Friedman H, Pross S, and Klein TW. 2006. Addictive drugs and their relationship with infectious diseases. FEMS *Immunol. Med. Microbiol.* 47 (3): 330-342.

Fritz KE, Woelk GB, Bassett MT, McFarland WC, Routh JA, Tobaiwa O, Stall RD. 2002. The association between alcohol use, sexual risk behavior, and HIV infection among men attending beer halls in Harare, Zimbabwe. *AIDS Behav.* 6 (3): 221-228.

Fritz K, Morojele N, and Kalichman S. 2010. Alcohol: the forgotten drug in HIV/AIDS. *Lancet* 376 (9739): 398-400.

Gerbi GB, Habtemariam T, Tameru B, Nganwa D, and Robnett V. 2009. The correlation between alcohol consumption and risky sexual behaviors among people living with HIV/*AIDS. J Subst. Use* 14 (2): 90-100.

Ghebremichael M, and Paintsil E. 2009. High Risk Behaviors and Sexually Transmitted Infections Among Men in Tanzania. *AIDS Behav.* 36(9):570-576.

Go VF, Srikrishnan AK, Parker CB, Salter M, Green AM, Sivaram S, Johnson SC, Latkin C, Davis W, Solomon S, Celentano DD. 2011. High prevalence of forced sex among non-brothel based, wine shop centered sex workers in Chennai, India. AIDS Behav. 15(1):163-71.

Golin C, Marks G, Wright J, Gerkovich M, Tien HC, Patel SN, Gardner L, O'Daniels C, Wilson TE, Thrun M, Thompson M, Raffanti S, and Quinlivan EB. 2009. Psychosocial characteristics and sexual behaviors of people in care for HIV infection: an examination of men who have sex with men, heterosexual men and women. *AIDS Behav.* 13 (6): 1129-1142.

Gmel G, Shield KD, and Rehm J. 2011. Developing a method to derive alcohol-attributable fractions for HIV/AIDS mortality based on alcohol's impact on adherence to antiretroviral medication. *Popul. Health Metr.* 9 (1): 5.

Gupta K, Vaidehi Y, and Majumder N. 2010. Spatial mobility, alcohol use, sexual behavior and sexual health among males in India. *AIDS Behav.* 14 Suppl 1: S18-S30.

Habbema JD, Van der Ploeg CP, Van Vliet C, Ndinya-Achola JO, and Fransen L. 1996. STDSIM: a decision support model for control of HIV and other STDs. Paper presented at *International Conference on AIDS 1996* .

Hargreaves JR, Morison LA, Chege J, Rutenburg N, Kahindo M, Weiss HA, Hayes R, Buvé A. 2002. Socioeconomic status and risk of HIV infection in an urban population in Kenya. *Trop. Med. Int. Health* 7 (9): 793-802.

Havanon N, Bennett A, and Knodel J. 1993. Sexual networking in provincial Thailand. *Stud. Fam. Plann.* 24 (1): 1-17.

Hendershot CS, Stoner SA, Pantalone DW, and Simoni JM. 2009. Alcohol use and antiretroviral adherence: review and meta-analysis. *J Acquir. Immune. Defic. Syndr.* 52 (2): 180-202.

Hill AB. 1965.The environment and disease: association or causation? *Proc R Soc Med.* 58:295–300.

Hou SI. 2009. HIV-related behaviors among black students attending Historically Black Colleges and Universities (HBCUs) versus white students attending a traditionally white institution (TWI). AIDS Care 21 (8): 1050-1057.

Jenness SM, Neaigus A, Hagan H, Murrill CS, and Wendel T. 2010. Heterosexual HIV and sexual partnerships between injection drug users and noninjection drug users. AIDS Patient Care STDS 24 (3): 175-181.

Jenness SM, Begier EM, Neaigus A, Murrill CS, Wendel T, and Hagan H. 2011. Unprotected anal intercourse and sexually transmitted diseases in high-risk heterosexual women. Am. J Public Health 101 (4): 745-750.

Jewkes RK, Dunkle K, Nduna M, Jama PN, and Puren A. 2010a. Associations between childhood adversity and depression, substance abuse and HIV and HSV2 incident infections in rural South African youth. Child Abuse Negl. 34 (11): 833-841.

Jewkes RK, Dunkle K, Nduna M, and Shai N. 2010b. Intimate partner violence, relationship power inequity, and incidence of HIV infection in young women in South Africa: a cohort study. *Lancet* 376 (9734): 41-48.

Johnson LF, Dorrington R, Bradshaw D, Pillay-Van Wyk V, and Rehle TM. 20 A.D. Sexual behaviour patterns in South Africa and their association with the spread of HIV: insight from a mathematical model. *Demographic Research* 21: 289-340.

Kalichman SC, Simbayi LC, Jooste S, Cain D, and Cherry C. 2006. Sensation seeking, alcohol use, and sexual behaviors among sexually transmitted infection clinic patients in Cape Town, South Africa. *Psychol. Addict. Behav.* 20 (3): 298-304.

Kalichman SC, Simbayi LC, Kaufman M, Cain D, and Jooste S. 2007. Alcohol use and sexual risks for HIV/AIDS in sub-Saharan Africa: systematic review of empirical findings. *Prev. Sci.* 8 (2): 141-151.

Kalichman SC, Simbayi LC, Vermaak R, Jooste S and Cain D.2008. HIV/AIDS Risks among Men and Women Who Drink at Informal Alcohol Serving Establishments (Shebeens) in Cape Town, South Africa. PREVENTION SCIENCE 9(1): 55-62.

Kalichman SC, Simbayi LC, and Cain D. 2010. HIV transmission risk behaviours among HIV seropositive sexually transmitted infection clinic patients in Cape Town, South Africa. *Eur. J Public Health* 20 (2): 202-206.

Kapiga SH, Lyamuya EF, Lwihula GK, and Hunter DJ. 1998. The incidence of HIV infection among women using family planning methods in Dar es Salaam, Tanzania. *AIDS* 12 (1): 75-84.

Karnell AP, Cupp PK, Zimmerman RS, Feist-Price S, and Bennie T. 2006. Efficacy of an American alcohol and HIV prevention curriculum adapted for use in South Africa: results of a pilot study in five township schools. *AIDS Educ. Prev.* 18 (4): 295-310.

Katz, I, Low-Beer, D. 2008. Why Has HIV Stabilized in South Africa, Yet Not Declined Further? Age and Sexual Behavior Patterns Among Youth. *Sexually Transmitted Diseases* 35 (10): 837-842.

Kelly JA, Murphy DA, Sikkema KJ, McAuliffe TL, Roffman RA, Solomon LJ, RA Winett, Kalichman SC and The Community HIV Prevention Research Collaborativ. 1997. Randomised, controlled, community-level HIV-prevention intervention for sexual-risk behaviour among homosexual men in US cities *The Lancet* 350 (9090): 1500-1505.

Kiene, SM, Simbayi, LC, Abrams A, Cloete, A. Tennen, H. and Fisher, JD. 2008. High rates of unprotected sex occurring among HIV-positive individuals in a daily diary study in South Africa: the role of alcohol use. *JAIDS - Journal of Acquired Immune Deficiency Syndromes.* 49(2): 219-226.

Koblin BA, Husnik MJ, Colfax G, Huang Y, Madison M, Mayer K, Barresi PJ, Coates TJ, Chesney MA, and Buchbinder S. 2006. Risk factors for HIV infection among men who have sex with men. *AIDS* 20 (5): 731-739.

Kumar GA, Dandona R, Kumar SG, and Dandona L. 2010. Behavioral surveillance of premarital sex among never married young adults in a high HIV prevalence district in India. *AIDS Behav.* 15 (1): 228-235.

Lambert G, Cox J, Hottes TS, Tremblay C, Frigault LR, Alary M, Otis J, and Remis RS. 2011. Correlates of Unprotected Anal Sex at Last Sexual Episode: Analysis from a Surveillance Study of Men who have Sex with Men in Montreal. *AIDS Behav.* 15 (3): 584-595.

Lane T, Raymond HF, Dladla S, Rasethe J, Struthers H, McFarland W, and McIntyre J. 2011. High HIV Prevalence Among Men Who have Sex with Men in Soweto, South Africa: Results from the Soweto Men's Study. *AIDS Behav.* 15 (3): 626-634.

Lauby JL, Semaan S, O'Connell A, Person B, and Vogel A. 2001. Factors related to self-efficacy for use of condoms and birth control among women at risk for HIV infection. *Women Health* 34 (3): 71-91.

Leigh BC, Morrison DM, Hoppe MJ, Beadnell B, and Gillmore MR. 2008. Retrospective assessment of the association between drinking and condom use. *J Stud. Alcohol Drugs* 69 (5): 773-776.

Lin D, Li X, Yang H, Fang X, Stanton B, Chen X, Abbey A, and Liu H. 2005. Alcohol intoxication and sexual risk behaviors among rural-to-urban migrants in China. *Drug Alcohol Depend.* 79 (1): 103-112.

London L. 2000. Alcohol consumption amongst South African farm workers: a challenge for post-apartheid health sector transformation. *Drug and Alcohol Depend.* 59: 199-206.

Lurie M. 2010. Population movement and the spread of HIV in southern Africa. In HIV / *AIDS in South Africa*, eds. SS Abdool Karim and Q Abdool Karim, 342-356. Cambridge University Press.

Mackesy-Amiti ME, Fendrich M, and Johnson TP. 2010. Symptoms of substance dependence and risky sexual behavior in a probability sample of HIV-negative men who have sex with men in Chicago. *Drug Alcohol Depend.* 110 (1-2): 38-43.

MacQueen KM, Nopkesorn T, Sweat MD, Sawaengdee Y, Mastro TD, and Weniger BG. 1996. Alcohol consumption, brothel attendance, and condom use: normative expectations among Thai military conscripts. *Med. Anthropol.* Q. 10 (3): 402-423.

Madhivanan P, Hernandez A, Gogate A, Stein E, Gregorich S, Setia M, Kumta S, Ekstrand M, Mathur M, Jerajani H, and Lindan CP. 2005. Alcohol use by men is a risk factor for the acquisition of sexually transmitted infections and human immunodeficiency virus from female sex workers in Mumbai, India. *Sex Transm Dis.* 32 (11): 685-690.

Madu SN, and Matla MP. 2003. Correlations for perceived family environmental factors with substance use among adolescents in South Africa. *Psychol. Rep.* 92 (2): 403-415.

Matos TD, Robles RR, Sahai H, Colon HM, Reyes JC, Marrero CA, Calderon JM, and Shepard EW. 2004. HIV risk behaviors and alcohol intoxication among injection drug users in Puerto Rico. Drug Alcohol Depend. 76 (3): 229-234.

McGrath N, Nyirenda M, Hosegood V, and Newell ML. 2009. Age at first sex in rural South Africa. *Sex Transm Infect.* 85 Suppl 1: i49-i55.

Mendoza-Perez JC, and Ortiz-Hernandez L. 2009. [Factors associated with infrequent condom use amongst men having sex with other men in Ciudad Juarez]. *Rev. Salud Publica* (Bogota.) 11 (5): 700-712.

Morisky DE, Chiao C, Ksobiech K, and Malow RM. 2010. Reducing alcohol use, sex risk behaviors, and sexually transmitted infections among Filipina female bar workers: effects of an ecological intervention. *J Prev. Interv.* Community 38 (2): 104-117.

Morojele NK, Kachieng'a MA, Mokoko E, Nkoko MA, Parry CD, Nkowane AM, Moshia KM, and Saxena S. 2006a. Alcohol use and sexual behaviour among risky drinkers and bar and shebeen patrons in Gauteng province, South Africa. *Soc. Sci. Med.* 62 (1): 217-227.

Morojele NK, Brook JS, and Kachieng'a MA. 2006b. Perceptions of sexual risk behaviours and substance abuse among adolescents in South Africa: a qualitative investigation. *AIDS Care* 18 (3): 215-219.

Morris M. 2010. Barking up the wrong evidence tree. Comment on Lurie & Rosenthal, "Concurrent partnerships as a driver of the HIV epidemic in sub-Saharan Africa? The evidence is limited". *AIDS Behav.* 14 (1): 31-33.

Morrison DM, Gillmore MR, Hoppe MJ, Gaylord J, Leigh BC, and Rainey D. 2003. Adolescent drinking and sex: findings from a daily diary study. *Perspect. Sex Reprod. Health* 35 (4): 162-168.

Murphy DA, Brecht ML, Herbeck DM, and Huang D. 2009. Trajectories of HIV risk behavior from age 15 to 25 in the national longitudinal survey of youth sample. *J Youth. Adolesc.* 38 (9): 1226-1239.

Nayak MB, Korcha RA, and Benegal V. 2010. Alcohol use, mental health, and HIV-related risk behaviors among adult men in Karnataka. *AIDS Behav.* 14 Suppl 1: S61-S73.

Nemoto T, Iwamoto M, Wong S, Le MN, and Operario D. 2003. Social factors related to risk for violence and sexually transmitted infections/HIV among Asian massage parlor workers in San Francisco. *AIDS Behav.* 8 (4): 475-483.

Nemoto T, Iwamoto M, Colby D, Witt S, Pishori A, Le MN, Vinh DT, and Giang IT. 2008. HIV-related risk behaviors among female sex workers in Ho Chi Minh City, Vietnam. *AIDS Educ. Prev.* 20 (5): 435-453.

Ngom P and Clark S. 2003. Adult mortality in the era of HIV/aids: Sub-Saharan Africa. *Workshop on hiv/aids and adult mortality in developing countries.* New York, 8-13 September 2003.

Nkansah-Amankra S, Diedhiou A, Agbanu HL, Toma-Drane M, and Dhawan A. 2010. Evaluating correlates of adolescent physical activity duration towards National

Health Objectives: analysis of the Colorado Youth Risk Behavioral Survey, 2005*. J Public Health (Oxf).*

Obot IS. 2007. Nigeria: alcohol and society today. *Addiction* 102 (4): 519-522.

O'Leary A, Broadwell SD, Yao P, and Hasin D. 2006. Major depression, alcohol and drug use disorders do not appear to account for the sexually transmitted disease and HIV epidemics in the southern United States. *Sex Transm Dis.* 33 (7 Suppl): S70-S77.

Operario D, Smith CD, Arnold E, and Kegeles S. 2011. Sexual risk and substance use behaviors among african american men who have sex with men and women. *AIDS Behav.* 15 (3): 576-583.

Parry CD, Pluddemann A, Steyn K, Bradshaw D, Norman R, and Laubscher R. 2005. Alcohol use in South Africa: findings from the first Demographic and Health Survey (1998). *J Stud. Alcohol* 66 (1): 91-97.

Parry C, Petersen P, Carney T, Dewing S, and Needle R. 2008a. Rapid assessment of drug use and sexual HIV risk patterns among vulnerable drug-using populations in Cape Town, Durban and Pretoria, South Africa. SAHARA. J 5 (3): 113-119.

Parry C, Petersen P, Dewing S, Carney T, Needle R, Kroeger K and Treger L. 2008b.Rapid assessment of drug-related HIV risk among men who have sex with men in three South African cities. *Drug Alcohol Depend.*1 95(1-2): 45-53.

Parry C, Rehm J, Poznyak V, Room R. 2009. Alcohol and infectious diseases: an overlooked causal linkage? *Addiction* 104 (3): 331–332.

Parry CD, Petersen P, Carney T, and Needle R. 2010a. Opportunities for enhancing and integrating HIV and drug services for drug using vulnerable populations in South Africa. Int. J Drug Policy 21 (4): 289-295.

Parry C, Rehm J, and Morojele NK. 2010b. Is there a causal relationship between alcohol and HIV? Implications for policy, practice and future research. African Journal of Drug and Alcohol Studies 9 (2).

Pattanaphesaj J and Teerawattananon Y. 2010. Reviewing the evidence on effectiveness and cost-effectiveness of HIV prevention strategies in Thailand. BMC. Public Health 10: 401.

Phillips CV and Goodman KJ. 2004. The missed lessons of Sir Austin Bradford Hill. Epidemiol. Perspect. Innov. 1 (1): 3.

Phorano O, Nthomang K, and Ntseane D. 2005. Alcohol abuse, gender-based violence and HIV/AIDS in Botswana: establishing the link based on empirical evidence. SAHARA. J 2 (1): 188-202.

Pithey A. and Parry C. 2009. Descriptive systematic review of Sub-Saharan African studies on the association between alcohol use and HIV infection. *Journal of Social Aspects of HIV/AIDS* 6(4): 155-169.

Poudel KC, Poudel-Tandukar K, Yasuoka J, Joshi AB, and Jimba M. 2010. Correlates of sharing injection equipment among male injecting drug users in Kathmandu, Nepal. *Int. J Drug Polic*y 21 (6): 507-510.

Prestage G, Grierson J, Bradley J, Hurley M, and Hudson J. 2009. The role of drugs during group sex among gay men in Australia. *Sex Health* 6 (4): 310-317.

Raj A, Reed E, Santana MC, Walley AY, Welles SL, Horsburgh CR, Flores SA, and Silverman JG. 2009a. The associations of binge alcohol use with HIV/STI risk and diagnosis among heterosexual African American men. Drug Alcohol Depend. 101 (1-2): 101-106.

Raj A, Cheng DM, Krupitsky EM, Levenson S, Egorova VY, Meli S, Zvartau EE, and Samet JH. 2009b. Binge drinking and unsafe sex: a study of narcology hospital patients from St. Petersburg, Russia. *Subst. Abus.* 30 (3): 213-222.

Ramjee G, and Gouws E. 2002. Prevalence of HIV among truck drivers visiting sex workers in KwaZulu-Natal, South Africa. Sex Transm Dis 29 (1): 44-49.

Randolph ME, Torres H, Gore-Felton C, Lloyd B, and McGarvey EL. 2009. Alcohol use and sexual risk behavior among college students: understanding gender and ethnic differences. *Am. J Drug Alcohol Abuse* 35 (2): 80-84.

Rehm J, Rehn N, Room R, Monteiro M, Gmel G, Jernigan D, and Frick U. 2003. The global distribution of average volume of alcohol consumption and patterns of drinking. *Eur. Addict. Res.* 9 (4): 147-156.

Rehm J, Room R, Monteiro M, Gmel G, Graham K, Rehn N, Sempos CT, Frick.U., and Jernigan D. 2004. Alcohol use. In Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors., eds M. Ezzati, AD Lopez, A Rodgers and CJL Murray, 959-1108. Geneva: World Health Organisation.

Rehm J, Anderson P, Kanteres F, Parry CD, Samokhvalov AV, and Patra J. 2009. Alcohol, social development and infectious disease. Toronto: Centre for Addiction and Mental Health. ISBN: 978-1-77052-444-6

Reisner SL, Mimiaga MJ, Bland S, Skeer M, Cranston K, Isenberg D, Driscoll M, and Mayer KH. 2010. Problematic alcohol use and HIV risk among Black men who have sex with men in Massachusetts. *AIDS Care* 22 (5): 577-587.

Rhodes SD, Bischoff WE, Burnell JM, Whalley LE, Walkup MP, Vallejos QM, Quandt SA, Grzywacz JG, Chen H, and Arcury TA. 2010. HIV and sexually transmitted disease risk among male Hispanic/Latino migrant farmworkers in the Southeast: Findings from a pilot CBPR study. *Am. J Ind. Med.* 53 (10): 976-983.

Rodríguez DC, AK, Kumarasamy N, Krishnan G , Solomon D, Johnson S,C. K. Vasudevan CK, Solomon R . Ekstrand ML. 2010. Two Sides of the Same Story: Alcohol Use and HIV Risk Taking in South India. AIDS and Behavior Volume 14, Supplement 1, Pages 136-146.

Roerecke M, Obot IS, Patra J and Rehm J. 2008. VOLUME OF ALCOHOL CONSUMPTION, PATTERNS OF DRINKING AND BURDEN OF DISEASE IN SUB-SAHARAN AFRICA, 2002. *African Journal of Drug & Alcohol Studies* 7(1): 1-15.

Rogers SJ, Ying L, Xin YT, Fung K, and Kaufman J. 2002. Reaching and identifying the STD/HIV risk of sex workers in Beijing. *AIDS Educ. Prev.* 14 (3): 217-227.

Room R. 2008. What do the North American diary and other event studies teach us about the role of drinking in risky sexuality? Paper presented at *Alcohol & Infectious Diseases Technical Meeting,* 15-18 July 2008.Vineyard Hotel, Cape Town, South Africa.

Rosengard C, Clarke JG, DaSilva K, Hebert M, Rose J, and Stein MD. 2005. Correlates of partner-specific condom use intentions among incarcerated women in Rhode Island. *Perspect. Sex Reprod. Health* 37 (1): 32-38.

Ryan GW, Stern SA, Hilton L, Tucker JS, Kennedy DP, Golinelli D, and Wenzel SL. 2009. When, where, why and with whom homeless women engage in risky sexual behaviors: A framework for understanding complex and varied decision-making processes. *Sex Roles* 61 (7-8): 536-553.

Säll L, Salamon E, Allgulander C, and Owe-Larsson B. 2009. Psychiatric symptoms and disorders in HIV infected mine workers in South Africa. A retrospective descriptive study of acute first admissions. *Afr. J Psychiatry* (Johannesburg) 12 (3): 206-212.

Samet JH, Pace CA, Cheng DM, Coleman S, Bridden C, Pardesi M, Saggurti N, and Raj A. 2010. Alcohol use and sex risk behaviors among HIV-infected female sex workers (FSWs) and HIV-infected male clients of FSWs in India. *AIDS Behav.* 14 Suppl 1: S74-S83.

Sander PM, Cole SR, Ostrow DG, Mehta SH, and Kirk GD. 2010. Determinants of alcohol consumption in HIV-uninfected injection drug users. *Drug Alcohol Depend.* 111 (1-2): 173-176.

Schensul JJ, Chandran D, Singh SK, Berg M, Singh S, and Gupta K. 2010. The use of qualitative comparative analysis for critical event research in Alcohol and HIV in Mumbai, India. AIDS Behav. 14 Suppl 1: S113-S125.

Schneider M, Norman R, Parry C, Bradshaw D, and Pluddemann A. 2007. Estimating the burden of disease attributable to alcohol use in South Africa in 2000. S. *Afr. Med. J.* 97 (8 Pt 2): 664-672.

Seedat M, Van Niekerk A, Jewkes R, Suffla S, Ratele K. 2009. Violence and injuries in South Africa: prioritizing an agenda for prevention. *Lancet* 374: 1011-1022.

Semple SJ, Strathdee SA, Zians J, and Patterson TL. 2010. Factors associated with sex in the context of methamphetamine use in different sexual venues among HIV-positive men who have sex with men. *BMC. Public Health* 10: 178.

Sheer VC. and Cline RW. 1995. Individual differences in sensation seeking and sexual behavior: Implications for communication intervention for HIV/AIDS prevention among college students. *Health Communication* 7 (3): 205-223.

Shisana O, Pehle T, Simbayi LC, Zuma K, Jooste S, Pillay-Van Wyk V., Mbelle N, Van Zyl J, Zungu NP, Pezi S, and & the SABSSM Implementation Team (2009). 2009. *South African National HIV prevalence, incidence, behaviour and communication survey, 2008: A turning tide among teenagers?* HSRC Press.

Shuper PA, Neuman M, Kanteres F, Baliunas D, Joharchi N, and Rehm J. 2010. Causal considerations on alcohol and HIV/AIDS--a systematic review. *Alcohol* 45 (2): 159-166.

Sikkema KJ, Watt MH, Meade CS, Ranby KW, Kalichman SC, Skinner D, and Pieterse D. 2011. Mental Health and HIV Sexual Risk Behavior among Patrons of Alcohol Serving Venues in Cape Town, South Africa. *J Acquir. Immune. Defic. Syndr.*

Simbayi, Leickness C.; Mwaba, Kelvin; Kalichman, Seth C. 2006. Perceptions of the combination of HIV/Aids and alcohol as a risk factor among STI clinic attenders in South Africa: Implications for HIV prevention. Social Behavior and Personality: an international journal, 34 (5): 535-544.

Singh K, Sambisa W, Munyati S, Chandiwana B, Chingono A, Monash R, and Weir S. 2010. Targeting HIV interventions for adolescent girls and young women in Southern Africa: use of the PLACE methodology in Hwange District, Zimbabwe. *AIDS Behav.* 14 (1): 200-208.

Singh SK, Schensul JJ, Gupta K, Maharana B, Kremelberg D, and Berg M. 2010. Determinants of alcohol use, risky sexual behavior and sexual health problems among men in low income communities of Mumbai, India. *AIDS Behav.* 14 Suppl 1: S48-S60.

Stein M, Herman DS, Trisvan E, Pirraglia P, Engler P, and Anderson BJ. 2005. Alcohol use and sexual risk behavior among human immunodeficiency virus-positive persons. *Alcohol Clin. Exp. Res.* 29 (5): 837-843.

Stein MD, Hanna L, Natarajan R, Clarke J, Marisi M, Sobota M, and Rich J. 2000. Alcohol use patterns predict high-risk HIV behaviors among active injection drug users. *J Subst. Abuse Treat.* 18 (4): 359-363.

Stockman JK, Campbell JC, and Celentano DD. 2010. Sexual violence and HIV risk behaviors among a nationally representative sample of heterosexual American women: the importance of sexual coercion. *J Acquir. Immune. Defic. Syndr.* 53 (1): 136-143.

Strathdee SA, Mausbach B, Lozada R, Staines-Orozco H, Semple SJ, Abramovitz D, Fraga-Vallejo M, Torre AL, Amaro H, Martinez-Mendizabal G, Magis-Rodriguez C, and Patterson TL. 2009. Predictors of sexual risk reduction among Mexican female sex workers enrolled in a behavioral intervention study. J Acquir. Immune. Defic. Syndr. 51 Suppl 1: S42-S46.

Tavares CM, Schor N, Franca I, Jr., and Diniz SG. 2009. Factors associated with sexual initiation and condom use among adolescents on Santiago Island, Cape Verde, West Africa. *Cad. Saude Publica* 25 (9): 1969-1980.

Tavarez MI, Chun H, and Anastario MP. 2011. Correlates of sexual risk behavior in sexually active male military personnel stationed along border-crossing zones in the Dominican Republic. *Am. J Mens. Health* 5 (1): 65-77.

The NIMH Multisite HIV/STD prevention Trial for African American Couples Group.(2010) The Contribution of Male and Female Partners' Substance Use to Sexual Risks and STDs Among African American HIV Serodiscordant Couples. *AIDS Behav.* 14 (5): 1045-1054.

Theall KP, Clark RA, Powell A, Smith H, and Kissinger P. 2007. Alcohol consumption, ART usage and high-risk sex among women infected with HIV. *AIDS Behav.* 11 (2): 205-215.

Thomas BE, Chandra S, Selvi KJ, Suriyanarayanan D, and Swaminathan S. 2009. Gender differences in sexual behaviour among people living with HIV in Chennai, India. *Indian J Med. Res.* 129 (6): 690-694.

Thomas RM, Plant MA, and Plant ML. 1990. Alcohol, AIDS risks and sex industry clients: results from a Scottish study. *Drug Alcohol Depend.* 26 (3): 265-269.

Todd CS, Nasir A, Stanekzai MR, Bautista CT, Botros BA, Scott PT, Strathdee SA, and Tjaden J. 2010. HIV, hepatitis B, and hepatitis C prevalence and associated risk behaviors among female sex workers in three Afghan cities. *AIDS* 24 Suppl 2: S69-S75.

Towe VL, Sifakis F, Gindi RM, Sherman SG, Flynn C, Hauck H, and Celentano DD. 2010. Prevalence of HIV infection and sexual risk behaviors among individuals having heterosexual sex in low income neighborhoods in Baltimore, MD: the BESURE study. *J Acquir. Immune. Defic. Syndr.* 53 (4): 522-528.

Townsend L, Rosenthal SR, Parry CD, Zembe Y, Mathews C, and Flisher AJ. 2010. Associations between alcohol misuse and risks for HIV infection among men who have multiple female sexual partners in Cape Town, South Africa. *AIDS Care* 22 (12): 1544-1554.

Townsend L, Jewkes R, Mathews C, Johnston LG, Flisher AJ, Zembe Y, and Chopra M. 2011. HIV risk behaviours and their relationship to intimate partner violence (IPV) among men who have multiple female sexual partners in Cape Town, South Africa. *AIDS Behav*. 15 (1): 132-141.

Tripathi A, Ruutel K, and Parker RD. 2009. HIV risk behaviour knowledge, substance use and unprotected sex in men who have sex with men in Tallinn, Estonia. *Euro. Surveill.* 14 (48).

Tsui HY., and Lau JT. 2010. Comparison of risk behaviors and socio-cultural profile of men who have sex with men survey respondents recruited via venues and the internet. *BMC. Public Health* 10: 232.

UNAIDS report on the global aids epidemic 2010 http://www.unaids.org/globalreport/global_report.htm

van Griensven F, Thienkrua W, Sukwicha W, Wimonsate W, Chaikummao S, Varangrat A, and Mock PA. 2010. Sex frequency and sex planning among men who have sex with men in Bangkok, Thailand: implications for pre- and post-exposure prophylaxis against HIV infection. *J Int. AIDS Soc.* 13: 13.

Verma RK, Saggurti N, Singh AK, Swain SN. 2010. Alcohol and sexual risk behavior among migrant female sex workers and male workers in districts with high in-migration from four high HIV prevalence states in India. AIDS Behav. 14 Suppl 1:S31-9.

Wang B, Li X, Stanton B, Zhang L, and Fang X. 2010. Alcohol use, unprotected sex, and sexually transmitted infections among female sex workers in China. *Sex Transm Dis.* 37 (10): 629-636.

Wechsberg WM, Luseno WK, Lam WK. 2005. Violence against substance-abusing South African sex workers: intersection with culture and HIV risk. AIDS Care 17 Suppl 1:S55-64.

Wechsberg WM, Luseno WK, Karg RS, Young S, Rodman N, Myers B, and Parry CD. 2008. Alcohol, cannabis, and methamphetamine use and other risk behaviours among Black and Coloured South African women: a small randomized trial in the Western Cape. *Int. J Drug Policy* 19 (2): 130-139.

Wechsberg WM, Wu LT, Zule WA, Parry CD, Browne FA, Luseno WK, Kline T, and Gentry A. 2009. Substance abuse, treatment needs and access among female sex workers and non-sex workers in Pretoria, South Africa. *Subst. Abuse Treat. Prev. Policy* 4: 11.

Wells BE, Kelly BC, Golub SA, Grov C, and Parsons JT. 2010. Patterns of alcohol consumption and sexual behavior among young adults in nightclubs. *Am. J Drug Alcohol Abuse* 36 (1): 39-45.

Wilson KS, Eggleston E, Diaz-Olavarrieta C, and Garcia SG. 2010. HIV/STI risk among male Mexican immigrants in Dallas, Texas: findings from a pilot study. *J Immigr. Minor. Health* 12 (6): 947-951.

Witte SS, Batsukh A, and Chang M. 2010. Sexual risk behaviors, alcohol abuse, and intimate partner violence among sex workers in Mongolia: implications for HIV prevention intervention development. *J Prev. Interv. Community* 38 (2): 89-103.

Wong FY, Huang ZJ, DiGangi JA, Thompson EE, and Smith BD. 2008. Gender differences in intimate partner violence on substance abuse, sexual risks, and depression among a sample of South Africans in Cape Town, South Africa. *AIDS Educ. Prev.* 20 (1): 56-64.

Woolf-King SE, and Maisto SA. 2011. Alcohol Use and High-Risk Sexual Behavior in Sub-Saharan Africa: A Narrative Review. *Arch Sex Behav* 40: 17–42.

Xiao Z, Palmgreen P, Zimmerman R, and Noar S. 2010. Adapting and applying a multiple domain model of condom use to Chinese college students. *AIDS Care* 22 (3): 332-338.

Zuma K, Gouws E, Williams B, and Lurie M. 2003. Risk factors for HIV infection among women in Carletonville, South Africa: migration, demography and sexually transmitted diseases. *Int. J STD AIDS* 14 (12): 814-817.

Section 3

Emerging Methods

14

Challenges in Healthcare in Multi-Ethnic Societies: Communication as a Barrier to Achieving Health Equity

Emine Kale and Bernadette Nirmal Kumar
Norwegian Centre for Minority Health Research, Oslo University Hospital
Norway

1. Introduction

The main learning objectives of this chapter are to gain knowledge and a better understanding of:

- Health and migration
- Migration and the challenges for healthcare services
- Equity in healthcare services
- Barriers to communication
- The use of interpreters in healthcare

2. Health and migration

Over 200 million people celebrated their last birthday outside their country of birth, characterizing the sheer scale, scope and extent of migration. The international migrant population (those living outside their country of birth) in 2010 was estimated at 214 million (3.1% of the global population), whereas UNDP has estimated that 740 million migrated within their country of birth. 75% of all international migrants are in 12% of all countries (IOM, 2002). Migration, globalization intensified by rapid transportation and communication technologies, and trade and commerce have all contributed to the evolving multicultural societies in nations all over the world. Modern day societies are characterized as being heterogeneous with increasing complexities of the heterogeneity.

As a part of this global development, West European countries, including Scandinavian countries, have become increasingly multicultural over the last few decades. In Norway half a million immigrants account for 12.2 % of the total population. In Oslo, the immigrant population stands at 28 %, the highest proportion in Norway (Statistics Norway, 2012). A consequence of these demographic changes is the challenge host societies face in integrating non-Western immigrant groups into existing healthcare services; the language barrier is the primary challenge for meeting the healthcare needs of the immigrant population.

3. Challenges in healthcare in multi-ethnic societies

Multicultural societies are often societies in transition that might be undergoing enormous changes. However, the changes are far from unidirectional, being positive and/or negative.

When migrants and their offspring are compared with other groups, changing socio-economic circumstances within and between generations in different migrant and ethnic groups can be linked to changing health patterns (Macbeth, 2001). This suggests that the health of adults might be related to exposures across their life course (Ben-Shlomo, 2004). The circumstances of migration and the social and health characteristics of resettlement are also key determinants of health (Wolffers, 2003). Post-migration determinants of health include the type of work migrants are expected to perform in the host country, the physical and housing conditions available to them, their language skills, remaining in contact with family, acquiring a new social network and their access to health and social services (Carballo M, 2004).

A recent epidemiological study in Norway, "The Oslo Immigrant Health Study", documents the health of migrants at the population level for the first time and indicates that the health of immigrant groups differs greatly compared to Norwegians and even more so with each other (Kumar, 2008). Only 30% of immigrant women from Pakistan and Turkey in the 59/60 year age group reported good health (Kumar, 2008). In all ethnic groups those with the highest education reported good health more frequently than others. Women reported more musculoskeletal disorders than men. However, men had higher proportions of myocardial infarction and stroke. Immigrant groups report mental distress more often than Norwegians, especially immigrant women (Kumar, 2008f) General obesity is a challenge for Turkish and Pakistani women in Oslo, as around 50% were obese (Body Mass Index >30). This is far higher than any of the other ethnic/gender groups. On the other hand, we found Vietnamese men and women with almost no obesity (3-4%).

Children of migrants are often caught at the crossroads between the majority (host) and minority (immigrant) cultures. Ethnic adolescents, therefore, land in **double jeopardy** with persisting unhealthy habits from their minority cultures and acquiring unhealthy habits from the majority as well. This is well illustrated in the case of boys from the Indian subcontinent in Oslo with high consumption of both full fat milk and carbonated soft drinks (Kumar, 2004f). Often the ability of children of migrants to adapt and adopt the host language and culture creates a perceived gap between them and their parents. Their immigrant parents fear that they are distancing themselves from their native values and behavioural patterns. The intra familial stress and parent-child conflicts may be precursors to low self-esteem, feelings of guilt and psychosocial morbidity among children of migrants (Kumar, 2010).

The poor health of immigrants is also reflected in their frequent use of health services. In the Oslo Immigrant Health Study, immigrants made a greater number of visits to the general practitioner (GP) and specialists compared to Norwegians. Turkish and Iranians visited the psychiatrist/psychologist most frequently. Emergency services were used most frequently by those from Turkey and least by the Norwegians (Kumar, 2008). Increased use of healthcare services may reflect: higher prevalence of mental distress related to lifestyle conditions among immigrants, reasons other than health problems cited when using health services or that their need for satisfactory healthcare is not met (FHI 2008). Particular challenges for migrants, such as language barriers, might be a contributing factor to ineffective communication and the increased use of healthcare services (Schyve, 2007).

Further analysis of the situation is recommended to gain a better understanding of the causes of this situation and to devise strategies to cope with it.

4. Equity in health and healthcare

4.1 The concepts of equity

Equity and equality are two concepts that are closely related, but not one and the same. While equality is well-defined, easily understood and measured, equity is not. Whitehead's definition of inequity refers to differences in health that are unnecessary and avoidable, and unfair and unjust (1985). While situations defined as unfair and unjust will vary depending on the place and time one has to examine the cause and judge the unfairness of the situation within the context of that particular society. Equity in health thus means that *every individual has a fair chance to attain their full health potential and, that no one should be disadvantaged from achieving this potential, if it can be avoided (p. 7)*

Therefore, the aim of policy for equity in health is to reduce or eliminate those health differences which result from factors which are considered to be both avoidable and unfair by creating equal opportunities for health and bringing health differentials down to the lowest level possible (Whitehead, 1985).

Braveman and Gruskin define equity as *the absence of disparities in health (or in the major social determinants of health) between social groups who have different levels of underlying social advantage/disadvantage - that is, different positions in the social hierarchy* (P. 254 2003). Populations who are already socially disadvantaged or socially excluded (for example, persons who are unemployed, homeless or members of a disenfranchised racial, ethnic or religious group) are at a further disadvantage with respect to their health. Social disadvantage here refers to two important attributes: the lower social hierarchy and less economically privileged position (or deprivation which can be relative or absolute). Equity has an ethical aspect which is based on the principle of distributive justice and linked to human rights, and can be assessed by comparing health and its social determinants among different more or less advantaged social groups.

4.2 Equity in healthcare

In practice equity means equal access to care, equal utilization for equal need and equal quality of care for all (Whitehead, 1985). Equal access to available care implies equal rights to the available services for everyone and a fair distribution of resources based on healthcare needs. Inequities in access arise when resources and facilities are unevenly distributed, for instance, greater availability in urban areas compared to scarce availability in deprived and rural areas. "Inconvenient" openings hour for clinics, communication barriers and large geographical distances and transport expenses can also be obstacles to equal access to health services.

While studies have documented over- or under-utilization of health services, further studies are required to understand better why the utilization rates are different. The variations in the utilization of services are not only indicators of inequities, but also related to the quality of services.

Equal quality of care for all means that providers must ensure that everyone gets the same high standard of professional care. However, when healthcare providers do not offer the same standards and quality of care to all individuals regardless of age, gender, religious belief or ethnic background, inequities will inevitably arise (Whitehead, 1985). Studies illustrating this phenomenon document negative health outcomes for the children of Spanish-speaking Latinos because of language barriers (Clemans-Cope et al., 2007). Language barriers preventing equal access to healthcare will be dealt with in further detail later in this chapter.

4.3 Equity in health policy

Health policies must address social determinants of health, such as improving living conditions, unemployment and working conditions of underprivileged groups, in order to achieve equity in health and healthcare services. These policies need to acknowledge that some groups in society are more disadvantaged than others. Hence, they face greater restrictions in their lifestyle choices, e.g., inadequate income and/or lower levels of education, which in turn limits where and how people live. Therefore, policies should enable people to adopt healthier lifestyles by increasing access to healthier lifestyles. (Whitehead, 1985)

An important aspect of achieving equity is to ensure user participation, involvement and empowerment, and to avoid a top down approach. This means that administrators need to make efforts to ensure that information is accessible to all, thereby making it easier for users to participate and influence the decision-making process. An important prerequisite in developing and improving equity is providing evidence by identifying the needs of different vulnerable groups. Research is also needed to monitor and evaluate the effectiveness of policies. An intersectoral approach is recommended as the determinants of inequities may be inherent in many other sectors in society (Shaw et al., 2006; Whitehead, 1985).

> *"What the scalpel is to the surgeon, words are to the clinicians…*
> *The conversation between doctor and patient is the hearth of the practice of medicine."*
>
> Woloshin et al. 1995:72

5. Communication barriers preventing equity in healthcare

The increasing number of immigrants from economically less privileged parts of the world to both Western Europe and Northern America, and the diversity this represents is often perceived as a challenge to existing healthcare services, which were traditionally organized to cover the needs of the native population. One of the most important challenges in healthcare is to be able to meet the needs of patients from immigrant backgrounds who are not proficient in the language used in the host country.

It is not only language barriers which may represent a challenge in healthcare, but also cultural differences in the perception of health/sickness and the sickness role, experiences of illness, help-seeking behaviour and health literacy level. These differences should not go unnoticed or be under-recognized. A language difficulty is perhaps the easiest problem to detect because it is often the most obvious one. Even so, there are usually no common procedures for systematically assessing the need for language assistance and ensuring adequate help. The healthcare providers, who usually have little or no training in how to evaluate patients' language abilities and often have no clear procedures for how to follow up after facing language barriers, seem often to be left alone to make the decisions themselves (Kale et al., 2010a). Even when a professional interpreter is attained, communication problems can arise due to lack of knowledge and skills on the part of healthcare provider regarding how to work together with the interpreter for optimal communication (Kale et al., 2010b). Healthcare institutions have responsibilities to ensure competency and procedures in their organizations in order to be able to give optimal health services to diverse populations for equal access and quality care for all.

5.1 Health literacy

Along with language barriers to effective communication in healthcare there are other potential barriers. Inadequate health literacy of patients is one of the reasons for difficulties in communicating effectively the treatment procedures or prescriptions in consultations and non-compliance. Health literacy refers to a mismatch between the healthcare provider's level of communication and the patient's level of comprehension of the medical information given to them (Weiss, 2003). Health literacy is not the same as literacy and is described elsewhere as an individual's ability to read, understand and use basic health information and services to critically evaluate the information and make appropriate health decisions. Health literacy requires a group of abilities like reading, understanding, remembering the information obtained, analysing and decision-making skills. For instance, the ability to read and understand prescription instructions, understanding informational materials and brochures, filling out forms and so on.

We do not have any statistics in Scandinavia, but according to the first large scale assessment done in 2003 in the USA, 21% of the adult population have basic and 14% below basic health literacy, while the majority (53 %) had intermediate health literacy (Kutner et al., 2006). In this study White and Asian/Pacific Islander adults had higher average health literacy than Black, Hispanic American Indian/Alaska Native and Multiracial adults. It has been shown that patients with low health literacy have less awareness of preventive health knowledge, less knowledge of their medical status and self-care instructions compared to people who are more literate (Weiss, 2003). This is an especially crucial issue for immigrant patients who are in a vulnerable position because of difficulties in getting the information they need in a language they can comprehend. In fact, elderly people, people with low socioeconomic status, unemployed people, minority ethnic groups and individuals who have recently immigrated and do not speak the majority language or have the majority language as their second language are in the risk group for low health literacy (Weiss, 2003)

(Ref: http://talk.onevietnam.org/vietnamese-americans-and-health-literacy-ready/

5.2 Language barriers

There are many models of how human communication proceeds between people from the simplistic and mechanical transmission model where a message is sent to f.exp. verbally from the so called "encoder" or sender to the "decoder" or receiver, to a more complex linear model of the Sender-message-Channel-Receiver Model (SMCR) of communication. There are also models which use a constructivist point of view in which communication is perceived as a collaborative and joint activity where reciprocally in the interaction is emphasized (Clark, 2007). When analysing communication, regardless of which communication model one employs to conceptualize the communication between people, one can still think of plenty of reasons why barriers to effective communication between the patient and healthcare provider arise in addition to lack of a common language, for instance, message *overload* (when a person receives too many messages at the same time) and *message complexity and misunderstanding* due to language-related factors (semantic or syntactical) and different perspectives because of cultural differences between the participants (Montana et al., 2008).

Cultural barriers might arise in meetings with migrants depending on what patients bring to the consultations such as their personality, earlier experiences, habits, attitudes, beliefs and prejudices. To a large extent these characteristics are shaped both by the society they belong to and are currently part of and therefore will be an integral part of the ongoing communication and thus also influence the future course of action. It is also about what the healthcare provider has in terms of personal cultural baggage, e.g., the *professional culture* he/she belongs to, as well as the *organizational or institutional culture* in which the interactions are embedded (Helman, 2001).

There has been increasing awareness about that the attitudes of therapists from majority populations with regard to various minorities and how this may influence diagnostic practices and the quality of the treatment provided (IOM 2002, ACP 2010). These reports document that minority patients have less access to necessary health services than White patients from the majority. Relatively significant differences in the treatment of various diseases were pointed out, for example, with regard to cancer, cardiovascular diseases, diabetes and psychiatric disorders among minorities, to the disadvantage of these patients. It is assumed that several factors could explain these differences, including health service organizations and consultation practices, of which the clinical uncertainty of physicians, caused by limited information/time constraints and negative attitudes to minority patients, may have a decisive impact.

It is essential that healthcare providers and patients communicate effectively to ensure that patients get proper help. Effective communication is defined as communication that is *"comprehended by both participants; it is usually bidirectional between participants, and enables both participants to clarify the intended message"* (Schyve & The Joint Commission, 2007, p.360) and requires a vast repertoire of skills in interpersonal processing like listening, observing, speaking, analysing and evaluating, all of which enable collaboration and cooperation. *"In the absence of comprehension, effective communication does not occur; when effective communication is absent, the provision of health care proceeds only with errors, poor quality, and risks to patient safety"* (p.360).

Ideally communication should be in the same language to be able to communicate efficiently and when not possible with qualified language assistance (i.e., using professional interpreters). Doctors view culture, ethnicity and language difficulties as barriers to both effective physician-patient communication and a satisfying working alliance (Johnson et al., 2004; Laveist et al., 2002; Meeuwesen et al., 2006). Immigrant patients living in Scandinavia report similar experiences in the existing studies. They complain about not being understood because of language problems and cultural differences, not having enough time to explain their problem and the doctors not being interested in their worries and concerns (Grønseth, 2006; Nielsen, 2005)

For many years the communication between physicians and patients has been made a topic in studies from different perspectives (Zimmermann et al., 2007; Kale et al., 2011). In a review article based on selected peer-reviewed studies of communication between physicians from majority populations and patients from a non-Western background, Schouten & Meeuwesen (2006) conclude that the research results obtained to date indicate considerable communication problems. Research also shows that healthcare workers underestimate the negative impact of language barriers and underuse interpreter services (Bischoff et al., 2010; Kale et al., 2010a).

The negative consequences of language-based obstacles in intercultural communication in healthcare are documented to some extent, but far less than adequately and seldom systematically (Divi et al., 2007). In Scandinavia, the consequences of poor communication between healthcare professionals and minority patients have only to some extent been discussed. For instance Essen (2001) examined stillbirths among women from Somalia, Ethiopia and Eritrea and found a lower quality of prenatal care in immigrant populations compared with the majority population. She pointed out the infrequent use of interpreters in the delivery ward as one of the most important reasons for the reduced quality of care.

Divi et al. (2007) point exactly to this lack of interest researchers have shown in this issue by referring to what Johnstone and Kanitsaki said in their article from 2004: '*…there is a paucity of literature specifically addressing the critical relationship that exists between culture, language, and patient safety, and the particular risks that patients from minority racial, ethno-cultural, and language backgrounds face when being cared for by healthcare professionals who do not know about, share, or understand either their culture or language*'

In order to fill this knowledge gap to some extent, Divi et al. (2007) studied the type and frequency of adverse events experienced by patients with Limited English Proficiency (LEP) and English-speaking patients in six U.S. hospitals. The findings showed that LEP hospital patients are more likely than their English-speaking counterparts to experience adverse events that result in harm, and the severity of that harm is often greater. Among 251 adverse events involving patients with LEP, 130 (52%) were related to communication problems whereas 36% of adverse events with English–speaking patients were related to communication problems.

- Misunderstandings and difficulties in uncovering misunderstandings
- Problems with giving preventive health information and in getting informed consent
- Difficulties with involving patients in their treatment and decision-making
- Inadequate comprehension of diagnoses and treatment
- Increased risk of misdiagnosis - both over- or under-diagnosis
- Inappropriate treatment or lack of treatment
- Over or underuse of healthcare services
- Increased use of unnecessary diagnostic resources
- Less adherence
- Frustrations and less satisfaction on both sides

Box 1. Potential negative effects of language barriers in healthcare (Jacobs et al., 2004; Moreno & Morales, 2010; Flores, 2003; 2005; Ngo-Metzer, 2003)

5.3 The use of interpreter services in order to overcome language problems

There is general agreement that a desirable way to overcome language barriers is the use of a professional interpreter. The widespread practice of using non-professionals, family members or friends, or bilingual staff on an ad hoc basis as interpreters has been discouraged (Frederics, 1996; Jareg et al., 2006). This is because emotional ties between the patient and their family and friends can interfere with the interpretation. Furthermore, non-professional interpreters cannot be expected to have adequate knowledge of medical terminology or a good enough command of both languages needed in order to impart the correct information. The patient's right of confidentiality and privacy may be breached if the patient feels forced to accept the presence of a family member as interpreter. Further, their presence can inhibit discussions regarding sensitive issues such as domestic violence, sexual abuse, psychiatric illness and other sensitive health problems like sexually transmitted diseases (Flores, 2003; 2005). Moreover, the use of a patient's minor-aged children in planned or acute consultations would be especially unethical and professionally irresponsible, not least towards the child, with regard to the child's best interests. This can even be against the principles of the conventions on the Rights of Children (Jareg & Pettersen, 2006)

5.3.1 "We take what we have": A questionnaire-based survey about the use of interpreters in Oslo

This title is an answer from one of the healthcare providers which describes the situation quite well considering what they usually do when language assistance is encountered with their patients. Although communication and language barriers between healthcare workers and patients have recently received attention internationally, in Scandinavia few studies have documented what healthcare workers do when they encounter language obstacles, the expectations they have of the interpreters and their evaluation of their own needs. Therefore, a questionnaire-based survey using a cross-sectional design was conducted with healthcare providers working at hospitals in Oslo as participants during 2004-2005.

Even though the immigrant's right to have a professional interpreter in encounters with the public health sector can be considered weakly anchored in existing legislation in several countries, there has been an increased emphasis on patient rights and the legal strengthening of these rights in Scandinavia, as well as in other Western countries. For instance according to the Patient Rights Law in Norway (1999), the patient has the right to contribute to or facilitate the consultation with the healthcare worker and the right to contribute to the choice of the available examinations and treatment methods. In order to be able to contribute, it is stated that information should be adapted to the patient's individual conditions, such as age, maturity, experience, culture and language background. In addition, healthcare workers are to ensure, to the best of their ability, that the patient has understood the content and meaning of the information.

The fact that responsibility is placed on the healthcare workers to evaluate the need for a professional interpreter emphasizes the importance of investigating the associated factors and situations related to this evaluation process and describe the common practices among healthcare workers. With that aim a cross-sectional survey study was conducted using a structured questionnaire in Oslo a few years ago. The survey was distributed to all general practitioners (GPs) at the primary care clinics in the three city districts in Oslo that have the highest percentage of non-Western immigrants. In addition, healthcare professionals in three hospitals that offer specialized health services to these city districts were included.

Even though the response rate was low in this study some interesting tendencies were uncovered. The study, in parallel with earlier studies, indicated that professional interpreters were underutilized in the health sector, considering the frequency of the language barriers experienced by providers in the study. Further, the answers indicated that the use of interpreters as a working method was not sufficiently embedded in the healthcare services as a standardized and quality assured procedure. Therefore, the use of interpreters seemed to be somehow incidental and dependent on the health provider's own knowledge and initiative.

Responses indicated that situations where healthcare workers did not use interpreters, even though the patient's understanding of Norwegian was insufficient, occurred quite often (in 28.8% of cases with doctors and 41.5% of cases with nurses). Further 25.3% of respondents indicated that they had often conducted the first conversation with a patient without knowing whether the patient's understanding of Norwegian was adequate.

How do conversations take place when there is a language barrier and how does this affect the health outcomes? This survey did not give an answer to these questions. Well-known impacts when facing language barriers is the feeling of defeat, vulnerability and perhaps helplessness on the part of the patient, but also on the part of the health provider. From clinical experiences, it is known that patients can blame themselves for not being able to speak the majority language better and often feel ashamed about this. Sometimes they do not concede that they have language difficulties, instead pretending to understand.

In this study, a large percentage of the respondents answered that they often tried to communicate with the patient without an interpreter (33%) in acute situations. This is thought-provoking considering that such situations might involve life-threatening conditions. Further findings indicated that healthcare providers had a tendency to resort to solutions that are *most easily available*, for example, using family or friends as interpreters or trying to communicate with the patient in spite of language barriers. More than half of the physicians and nurses responded that they often communicated with the patient through family member(s) or friends. What were the reasons for not using professional language assistance? The healthcare workers stated often that the reasons for not arranging an interpreter were impracticality, it being too time-consuming and poor access to interpreter services.

Further, a considerable percentage of the survey participants expressed dissatisfaction with both their own methods of working with interpreters and with the interpreter's qualifications.

5.3.2 Implications of this study for healthcare services

One of the implications of this study is that the existing practices can have negative consequences for equal access to healthcare services for patients with limited majority language proficiency and inadequate health literacy. The Patient Rights Law in Norway places the responsibility on healthcare workers and healthcare institutions to guarantee the patient's right to information and input by providing optimal communication with patients. The healthcare providers, who usually have no training in how to evaluate patients' language abilities, seem often to be left alone to make the decisions. A newly conducted study indicates that healthcare providers and patients might evaluate quite differently whether or not a language barrier exists (Le C, 2011). Therefore, it is important that administrators at healthcare services and healthcare policy makers are aware of their responsibility to secure the knowledge base and procedures necessary to fulfil the intention of the laws.

Better routines and procedures in the workplace for the effective organization of work with interpreters and a higher awareness and competence at the institutional level about which measures should be taken in order to adapt healthcare services for patients with limited majority language proficiency and inadequate health literacy are recommended.

Before the consultation

- Prefer professional interpreters rather than ad hoc solutions with for example family members
- Set up time to meet the interpreter to explain the goal of the consultation
- Check if the interpreter's background characteristics and position in the community might affect the relationship between interpreter and the patient negatively
- Explain the field in which the interpretation is needed and some specifics aspects of the topic which the interpreter should be aware of and prepared for
- Inform the interpreter briefly about your methods and approaches
- Agree upon a cooperation model

During the consultation

- Sit in a triangle such that you face the patient and the interpreter is sitting beside both of you
- Ensure that the interpreter always explains his/her role and gives information about the confidentiality rule at the beginning of the consultation
- Look at the patient when the interpreter speaks
- Use direct speech (e.g., "you" instead of "tell her/him that…)"
- Avoid long sentences and jargon
- Be aware of signals from the patient and the interpreter about the quality of the communication and interaction
- To secure effective communication ask control question to the patient, for example, ask what he/she understood of the information/instructions you have just given
- Do not involve the interpreter as a cultural broker or mediator, unless he/she has a defined role as such
- Be aware that mediated communication is different than direct immediate communication, and has its own limits

After the consultation

- Give a debriefing to the interpreter if it has been a difficult consultation emotionally
- Evaluate the cooperation and give feedback
- Use the same interpreter if possible to assist in developing a professional working alliance being established between you and the interpreter, as well as between the patient and the interpreter - this can be especially important in mental health services

Box 2. Some recommendations for working with interpreters in medical settings

6. Conclusions

- With over 200 million migrants in the world today, steadily increasing migration is a key driver of multiethnic societies.
- Multi-ethnic societies multiply the challenges for healthcare and these range from varying health behaviours, beliefs and attitudes, diseases, communication, language and cultural barriers, requirements based on religion, lack of information, personal biases, stereotyped views, individual racism to institutional (health system) bias and enforcement of laws requiring equal opportunities in employment and other walks of public life.

- Unless there is a focus on health inequities disadvantaged groups will not have a fair opportunity to attain their full health potential. Health equity should not only be seen in the light of rights, laws and at the macro level of health systems, but also in access and quality of care and equal utilization for equal need.
- Communication in healthcare is influenced by many factors. Language barriers and inadequate health literacy can, among other things, influence communication negatively in healthcare encounters. It seems that the influence of these factors on communication and health outcomes are often underestimated by healthcare providers and policy makers.
- Existing legislation in several countries has increased emphasis on patient rights and the legal strengthening of these rights. This implies that healthcare workers and healthcare institutions have the responsibility to guarantee the patient's right to information by providing optimal communication with patients and communication should be adapted to patient needs.
- An optimal way to overcome language barriers is to ensure the assistance of professional interpreters, but studies mentioned above indicate that the decision-making of healthcare providers regarding whether or not professional language assistance is needed is often influenced by hectic working conditions, making the providers resort to solutions that are most easily available, but not necessarily optimal.
- These practices can have negative consequences for equal access to healthcare services for patients with limited majority language proficiency and inadequate health literacy.
- Finally, given that multi-ethnic societies are here to stay, further operational research and development, and implementation of good practices are critical to both tackling health inequities and overcoming barriers in communication.

7. References

American College of Physicians. Racial and Ethnic Disparities in Health Care, Updated 2010. Philadelphia: American College of Physicians; 2010: Policy Paper. (Available from American College of Physicians, 190 N. Independence Mall West, Philadelphia, PA 19106)

Baker D, Parker R, Williams M, Coates WC & Pitkin K. (2006). Use and effectiveness of interpreters in an emergency department. *J Amer Med Assoc*, 275:783–788.

Ben-Shlomo Y & Kuh D. (2004). *A Life Course Approach to Chronic Disease Epidemiology*. 2 ed. New York: Oxford University Press.

Bischoff A & Hudelson P. (2010). Communicating with foreign language-speaking patients: is access to professional interpreters enough? *J Travel Med*, 17:15–20.

Braveman P & Gruskin S. (2003). Defining equity in health. *J Epidemiol Community Health*. 57:254-258.

Carballo M, Smajkic A, Zeric D, Dzidowska M, Gebre-Medhin J & Van Halem, J. (2004). Mental health and coping in a war situation: the case of Bosnia and Herzegovina. *J Biosoc Sci*. Jul;36(4):463-77.

Clark, HH. (1996). *Using Language*. Cambridge University Press.UK.

Clemans-Cope & Kenney G. (2007). Low income parents' reports of communication problems with health care providers: effects of language and insurance. *Public Health Rep*. Mar-Apr; 122(2):206-16.

Divi C, Koss R G, Schmaltz SP & Loeb JM. (2007). Language Proficiency and Adverse Events in U.S. Hospitals: A Pilot Study, *Inter Journal for Quality in Healthcare,* April 19(2):60–67.

Essen B. (2001). *Prenatal morality among immigrants from Africa's Horn. The importance of experience, rationality, and tradition for risk assessment in pregnancy and childbirth.* Scientific treatise. Medical faculty, University of Lund, Malmø, Sweden.

FHI (2008). Norwegian Inst. for Public Health. *The Oslo immigrant profile*. Oslo.

Flores G, Laws MB, Mayo SJ, et al. (2003). Errors in medical interpretation and their potential clinical consequences in pediatric encounters. *Pediatrics*.111:6-14.

Flores G. (2005). The impact of medical interpreter services on the quality of health care: a systematic review. *Med Care Res Rev*. 62:255-299.

Frederics C. (1996). Using non-professional interpreters in a multiethnic primary care clinic. University of Malaya, Malaysia. (Accessed 24 January, 2005, at http://www.criticallink.org/ journalsc12/10.htm.).

Grønseth, AS. (2006). *Lost selves and lonely persons. Experiences of illness and well-being among Tamil refugees in Norway*. PhD dissertation at NTNU.

Helman CG. (2001). *Culture, Health and Illness*. (Fourth edition) Arnold, London.

International Organization for Migration (2010). *World Migration Report 2010 – the future of migration: Building capacities for change*. Geneva: IOM.

Institute of Medicine (IOM) (2002). *Unequal Treatment: Confronting racial and Ethnic Disparities in Healthcare*. Washington, DC: National Academies Press.

Jacobs E, Shepard D, Suaya J & Stone E. (2004). Overcoming language barriers in health-care: Costs and benefits of interpreter services. *American Journal of Public Health*.94:866–869.

Jareg K & Pettersen Z. (2006).*Tolk og tolkebruker – to sider av samme sak. [The interpreter and the interpreter's user - two aspects of the same issue]* Bergen/Oslo: Fagbokforlaget/NAKMI.

Johnson RL, Roter D, Powe NR & Cooper LA. (2004). Patient Race/Ethnicity and Quality of Patient–Physician Communication During Medical Visits. *American Journal of Public Health*: Vol. 94, No. 12, pp. 2084-2090.

Kale E & Syed HR. (2010 a). Language barriers and the use of interpreters in the public health services. A questionnaire-based survey. *Patient Educ Couns*. 81, 187-191.

Kale E, Ahlberg N & Duckert F. (2010 b). Hvordan håndterer helsepersonell språklige barrierer? En undersøkelse av tolkebruk i helsevesenet (How do healthcare providers deal with language barriers?) *Tidskrift for Norsk Psykologforening*, 47, 818-823.

Kale E, Finset A, Eikeland HL & Gulbrandsen P. (2011). Emotional Cues and Concerns in Hospital Encounters with Non-Western Immigrants as compared with Norwegians. An Exploratory Study. *Patient Educ Couns*, 84, 325-331.

Kazzi B, Cooper C. (2003). Barriers to the use of interpreters in emergency room pediatric consultations. *J Pediatric Child Health.39*:259–263.

Kumar, BN, G. Holmboe-Ottesen NL & Wandel M.(2004). Ethnic Differences in Body Mass Index and Associated Factors of Adolescents From Minorities in Oslo, Norway: A Cross-Sectional Study. *Public Health Nutrition* 7(8): 999–1008.

Kumar BN, Grøtvedt L, Meyer HE, Sogaard AJ, Strand BH. (2008). *The Oslo Immigrant Health Profile*. Oslo. Nasjonalt folkehelseinstitutt (7).

Kumar BN, Viken B. (2010). (Eds). *Folkehelse i et migrasjonsperspektiv*. Bergen: Fagbokforlaget.

Kutner M., Greenberg E., JinY., and Paulsen C. (2006). *The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy* (NCES 2006–483).U.S.

Department of Education. Washington, DC: National Center for Education Statistics.

Le, C. (2011). *Tolkebruk i helsetjenesten (Use of interpreters in health services)*, Master thesis. UIO.

Lov 1999-07-02 nr. 63: Lov om pasientrettigheter. [Act relating to the rights of patients]

Laveist TA & Nuru-Jeter A. (2002). Is doctor-patient race concordance associated with greater satisfaction with care? *J Health Soc Behav*. 4:296-306

Macbeth H & Shetty PS. (2001). *Health and ethnicity*. London: Taylor & Francis

Meeuwesen L, Harmsen JA, Bernsen RM, et al. (2006). Do Dutch doctors communicate differently with immigrant patients than with Dutch patients? *Soc Sci Med*; 63: 2407–17.

Meyer B. (1998). Interpreter-mediated doctor-patient communication: The performance of non-trained community interpreters. Seminar: Second international conference on interpreting in legal, health, and social service settings, Vancouver, Canada, 19-23. (Accessed 8 September, 2005, at http://www.criticallink.org/English/criticallink2_papers.htm.)

Montana P J. & Charon, B H. (2008). *Management*. 4th ed. New York. Barron's Educational Series, Inc.

Moreno G. & Morales LS. (2010). Hablamos Juntos (together we speak): Interpreters, provider communication, and satisfaction with care. *J Gen Intern Med.* Dec; 25(12):1282–8.

Morales LS, Cunningham WE, Brown JA, Liu H & Hays RD. (1999). Are Latinos less satisfied with communication by health-care providers? *J Gen Intern Med.* 14: 407-417.

Nielsen AS. (2005). *Smertelige erfaringer (Painfull experiences).* PhD. dissertation, København Universitet.

Ngo-Metzger Q, Massagli MP, Clarridge BR, Manocchia M et al. (2003). Linguistic and cultural barriers to care. *J Gen Intern Med*. 18:44–52.

Pöchhacker F. (2000). The Community Interpreter's Task: Self-Perception and Provider Views. R P.Roberts, S E.Carr, D.Abraham & A. Dufour A (Eds.), *The Critical Link 2: Interpreters in the community*. 49-65. John Benjamins Publishing

Schyve PM. (2007). Language Differences as a Barrier to Quality and Safety in Health Care: The Joint Commission Perspective. *J Gen Intern Med*. 2007 November; 22(Suppl 2): 360–361.

Shaw M, Dorling, D & Smith GD. (2006). Poverty, social exclusion and minorities. Marmot M & Wilkinson R G. (Eds), *Social determinants of Health.* 2. Ed. Oxford University Press, New York.

Weiss BD. (2003). *Health Literacy. A manual for clinicians*. American Medical Association Foundation and American Medical Association. Accessed 21. Januar 2012 http://www.acibademsaglik.com/upload/pdf/literatur40.pdf)

Whitehead M. (1992). The concepts and principles of equity in health. *Int J Health Serv* 22:429–445. (First published with the same title from: Copenhagen: World Health Organisation Regional Office for Europe,1990 [EUR/ICP/RPD 414])

Woloshin S, Schwartz LM, Katz SJ & Welch HG. (1997). Is language a barrier to the use of preventive services? *J Gen Intern Med*. 12:472-477.

Wolffers I, Verghis S, Marin M. (2003).Migration, human rights, and health. *Lancet*. Dec; 362(9400):2019-20.

Zimmermann C, Del Piccolo L, Finset A. (2007).Cues and concerns by patients in medical consultations: a literature review. *Psychol Bull* 133:438–63.

15

Public Health Research and Action: Reflections on Challenges and Possibilities of Community-Based Participatory Research

S. Lazarus[1,2,3], B. Duran[4], L. Caldwell[5,6,7] and S. Bulbulia[1]
[1]*Safety and Peace Promotion Research Unit (SAPPRU)*
[2]*Institute for Social and Health Sciences, University of South Africa*
[3]*Faculty of Education, University of the Western Cape*
[4]*Indigenous Wellness Research Institute (IWRI), University of Washington, Seattle*
[5]*University of Nebraska, Lincoln, Rhodes College*
[6]*University of Memphis*
[7]*Annie E. Casey Foundation*
[1,2,3]*South Africa*
[4,5,6,7]*USA*

1. Introduction

Research has an important role to play in public health, providing new knowledge which is generally used to inform application in all areas of health. In addition to the traditional role of research in health, where the focus is on applied knowledge production, research can also be used as a form of *action*, and therefore is a recognised area of activities which, in itself, can be a form of intervention (Springett & Wallerstein, 2008). As will be come clear below, there are various terms given to this type of research, and there are varied interpretations of what this kind of activity entails. The main focus, however, is generally on both *knowledge* production, *and action* in community or social affairs (Lazarus, 2007). Within the context of action or intervention research, the element of *participation* is sometimes included as a central principle and characteristic. The emphasis in participatory forms of research is on facilitating various levels of participation of the 'researched' in the study.

In this chapter, the focus is on the utilisation, possibilities and challenges of Community-Based Participatory Research (CBPR) as a tool for both research and action in public health. This chapter makes some important contributions to public health in general, and research more specifically. First, it emphasises the value of research as a public health intervention. Second, it emphasises the importance of participation in public health research. Third, it provides an important critical lens to public health research, drawing on social critical perspectives which highlight power dynamics in research endeavours. Fourth, it provides important insights from the authors' collective experiences of working in marginalised communities, in the USA (specifically American Indian/Alaska Native (AIAN), and African American contexts), and in South Africa.

The chapter commences with an overview of CBPR as a particular approach, and locates this approach within broader meta-theoretical paradigmatic frameworks. Drawing on both literature and personal experiences of the authors, the next section focuses on the utility, possibilities and potential outcomes of CBPR. Key challenges often faced in this kind of activity are then identified. The authors then draw on their own experiences to provide some suggestions for how some of these challenges can be addressed.

2. CBPR as an approach to research and action in public health

Community-engaged research is labeled in different ways, including the following *terms:* Community-Based Participatory Research (CBPR) (Israel, 2005; Minkler & Wallerstein, 2008), Participatory Action Research (PAR) (Cornwall & Jewkes, 1995), Action Research (Reason & Bradbury, 2001, 2008), and Research Practice Networks (Green & Hickner, 2006; Westfall, Mold & Fagnan, 2007), amongst others. Community engagement, a term currently commonly used in university circles in South Africa, is also often used in this context, although it goes beyond only research activities, including various levels of community participation.

The following *definitions* of CBPR are useful. Israel et al. (2005) state that CBPR refers to a partnership approach to research that equitably involves community members, organisation representatives, and researchers in all aspects of the research process. According to the Kellogg Foundation's Community Scholars Programme, CBPR is a collaborative approach to research that equitably involves all partners in the research process and recognizes the unique strengths that each brings. CBPR begins with a research topic of importance to the community with the aim of combining knowledge and action for social change to improve the community.

In contrast to an instrumental research approach, *CBPR* constitutes a worldview reflected through an applied approach or process which includes the following specific *characteristics*:

- The Community as the unit of identity, solutions and practice (Israel, Schulz, Parker & Becker, 2001; Schulz et al., 2002)
- Community engagement at all levels of the research, from problem identification and theory development to sustainability
- A multi-level focus of change, including individual, community and social determinants (Hawe, Shiell & Riley, 2009)
- Epistemological diversity in theory and methods; including indigenous decolonizing approaches, empowerment, feminist, queer, systems theory and other critical approaches (Wallerstein & Duran, 2008)
- Research that includes clearly outlined mutual benefits and co-learning between partners (Israel et al., 2001, Schulz et al., 2002)
- A long term process and commitment to sustainability (Israel, 2005; Minkler & Wallerstein, 2008)

This approach is in line with general *community engagement* approaches which include the following key characteristics (Attree & French, 2007; Herbertson, Ballesteros, Goodland & Munilla, 2009; Popay, 2006; Rifkin, Lewando-Hundt & Draper, 2000):

- Understanding the history and current dynamics of the community, and respecting its uniqueness
- Aligning community engagement with the community's own plans and agenda for community development

- Strengthening and sustaining communities
- Establishing appropriate structures and processes to elicit and represent community views, and ensuring that they are accountable to the community
- Working with and through a range of groups and organisations, not just formal channels
- Fostering participation of the community in all key steps

The community engagement approach aligns with CBPR in its focus on the relationship between the 'university and community', with CBPR emphasizing the need for *shared control*, thus emphasizing the power relations in the process.

Historically, the CBPR approach has been informed by action research, led by Kurt Lewin in the 1940s. Lewin's action research approach included cycles of action, reflection, problem-solving and decision-making for new actions, with organisational change being a major focus (Lewin, 1938, 1951, 1997; Lewin & University of Michigan Research Center for Group Dynamics, 1975). In the 1970s, participatory research became popular, as a result of radical critiques of research traditions by social scientists from Asia, Africa, and Latin America in particular. Orlando Fals-Borda, most notably, combined scientific research, adult education and political action that aimed to: (a) raise levels of consciousness (b) empower class and group interests to organize, and (c) evaluated by concrete benefits to communities, not abstractions (Fals-Borda & Rahman, 1991). This included critiques of structural underdevelopment and the need to redistribute inequitable structures, challenges to academic distance from communities, and the development of new academic discourses of feminism, post-colonialism, post structuralism (Wallerstein & Duran, 2008).

Much of the above is aligned with the principles of *empowerment* (Fetterman & Wandersman, 2005) where capacity building, community ownership and democratic participation, within a, social justice framework, are emphasized. This incorporates a key indigenous principle which states: "Don't plan about us, without us!" (NACCO, 2001).

It is important to note that the *principle of participation* is central to this approach and that not all collaboration meets standards of CBPR. One way to view community engagement is on a continuum from minimal participation to full participation of the community concerned. These different positions on the continuum can be seen as either consultation, involvement, or engagement (Hashagen, 2002).

- *Consultation* suggests simply providing information to the community and requesting feedback, and carries no undertaking that there is to be a shift in what is done *or* how it is done.
- *Involvement* carries a stronger message, implying that the academic institution decides on the structures and decision-making processes, and that the community needs to be encouraged to become involved in them. The community has no part in deciding on the suitability of those structures or processes.
- *Engagement* suggests a different sort of relationship which avoids making assumptions about communities, asks for a dialogue, and implies that the development of the relationship itself will need to be a focus of attention.

Community research in general draws on a variety of *research methodologies*, which can be viewed on a continuum of control over phenomena <------> collaboration Heller et al. (1984). Specific research methods used include: participant observation, ethnographic approach, network analysis, studies of community populations, and social indicators (measures of social and community well-being), action research, simulation, field experiments, quasi-experimental approaches, and time-series designs.

Some examples of different methods used in public health include: participatory community case studies, including random community trials (RCT) (Pokorny et al., 2004); epidemiological research including both science and community participation (Torres-Harding et al., 2004); participatory project-based research, where both scientific standards and maximum impact are considered important (Stoeker, 2005); participatory rural appraisal (including historical mapping and inventory asset mapping (Stoeker, 2005); participatory or empowerment evaluation linked to public health interventions, where "participatory evaluation is a health-promoting intervention in itself" (Springett and Wallerstein, 2008, p. 205); and community-driven asset identification, referred to as "barefoot epidemiology" by Minkler and Hancock (2008), which uses a variety of traditional and innovative methods. Community asset mapping is used as a tool for identifying community resources, promoting community pride and eliciting community-embedded knowledge (Minkler & Hancock, 2008).

Health promotion research, which usually includes some form of programme development, implementation and evaluation (Reddy et al., 2003), targets populations, communities or settings, and is oriented towards community change; draws on the strengths of the community; takes the sociocultural context into account; and is usually community-based, emphasizing "empowering people through mediating structures, networks, and community institutions" (Revenson & Schiaffino, 2000, p. 473).

CBPR could use any of the above methods, but the *process* is guided by the values and principles outlined above, with community participation and ownership being central. This approach to research is generally located within a *participatory paradigm* (Guba & Lincoln, 2005), but draws from a number of research paradigms such as post-positivist, constructivist and transformative perspectives (Mertens, 2005), including both hermeneutic and critical traditions (Springett & Wallerstein, 2008). The participatory paradigm reflects holistic, systemic, and relational worldviews (Bradbury & Reason, 2008; Guba & Lincoln, 2005), where knowledge is generated as partners (Springett & Wallerstein, 2008). The participatory paradigm outlined in Guba and Lincoln's (2005, p. 195-199) updated version of 'basic beliefs of alternative inquiry paradigms' focuses on the generation of a participative reality, co-created by mind and cosmos; a critical subjectivity, favouring experiential, propositional and practical knowing, and co-created findings; political participation in collaborative action inquiry, within a language grounded in shared experiential contexts; pursued in communities of practice; focusing on transforming the world in the service of human flourishing; including self-reflective action; drawing on voices through narrative, movement, song, dance and other presentational forms; and pursued by co-researchers who learn through active engagement in the process, using democratic qualities and skills.

3. Utility, possibilities and potential outcomes of CBPR

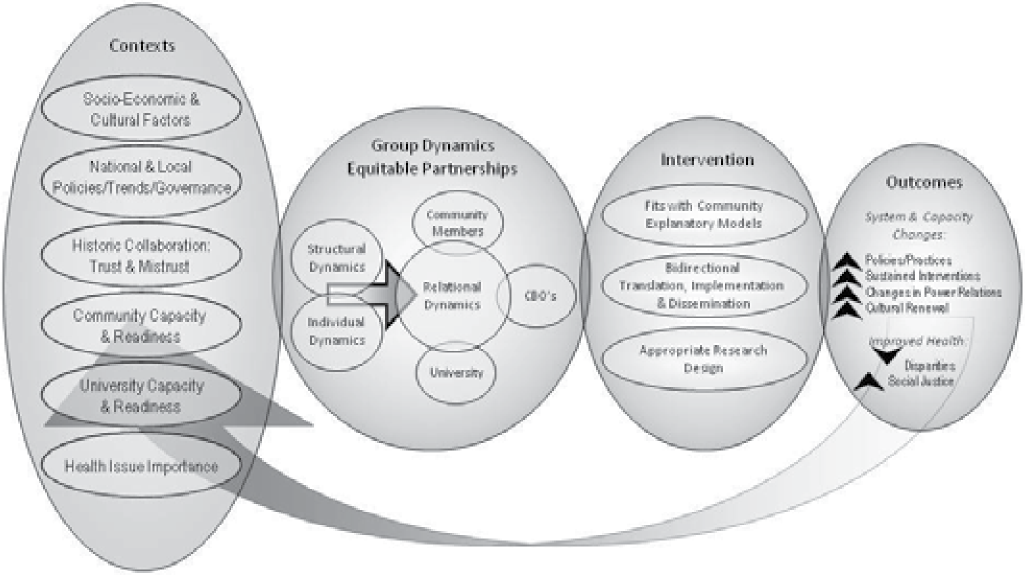
Why should we use a CBPR approach? Some of the reasons are briefly outlined below.

- There is a increasing interest in health disparities which are appropriately addressed by this approach.

- There is an increasing community and funder demand for community-driven research.
- There have been disappointing results in intervention research Complex health and social problems are ill-suited to "outside expert" research.
- It supports implementation and research dissemination.
- It supports the principle of democracy, and addresses issues of power and domination.
- It builds capacity and reduces dependency on "professional outsiders".
- It ensures cultural and local competence.
- It facilitates sustainability.
- It enhances fit and productivity of programmes.
- It produces more valid research.
- Ethical considerations are well addressed.

Minkler and Baden (2008) argue that CBPR often results in more effective research. It is externally valid in terms of utility, and internally valid because people are likely to be more honest and willing to participate, resulting in more accurate data. The development of local knowledge is facilitated, and CBPR methods contribute to better translation of research and practice by facilitating community-academic communication and flows of knowledge.

The *potential outcomes* of this approach are clear. As Wallerstein's (2008, 2010) *Interactive CBPR Conceptual Model* (refer Figure 1 below) outlines, at the system level, capacities are developed, policies and practices are improved, and sustained interventions can be developed. At the broadest level of meeting the goal of improving health and addressing disparities, this approach contributes to social justice.



Source: Wallerstein, 2008, p. 177; Wallerstein, 2010, p. 131

Fig. 1. Conceptual Logic Model of Community-Based Participatory Research: Processes to Outcomes

Since 2009, the Indigenous Wellness Research Institute, the University of New Mexico Center for Participatory Research, and the National Congress of American Indians Policy Research Center have collaborated as investigators to better understand how Community Based Participatory Research (CBPR) works to improve health and health equity. Their NIH project ("Research for Improved Health: A National Study of Community-Academic Partnerships," funded for 2009-2013) aims to (a) better understand CBPR practice variability across contexts, conditions and populations, including American Indian/Alaska Native communities, communities of colour and others which face health disparities; (b) identify promoters and inhibitors of CBPR partnership success; (c) better understand CBPR pathways and promising partnership practices that lead to improved health status; and (d) further develop appropriate research/evaluation measurement tools and methods to assess CBPR partnership effectiveness. The Interactive CBPR Conceptual Model which is being developed by this research team is available as a web tool, allowing users to download surveys (instruments) and individual items measures (variables), with associated information useful for evaluating CBPR partnerships and assessing partnership characteristics.

4. Key challenges of community-based participatory research

The following set of *challenges* linked to conducting research within a CBPR approach has been identified in the literature.

There are various *tensions* inherent in this approach to research. These tensions include (a) science <---> community participation, (b) science/research <---> practical goals/action, and (c) control over phenomena <---> collaboration continuum. These tensions are described in more detail below.

Issues of *control* are central to challenges experienced in CBPR. In particular, non-randomized designs, which are often favoured in this kind of research, raise a number of challenges in CBPR (Farquhar and Wing, 2008: Heller et al., 2004; Springett & Wallerstein, 2008; Stoeker, 2005). This means that (a) you cannot randomly assign people, resulting in selection bias, (b) the research often cannot be replicated as communities are different, (c) disparities in programmes means you cannot generalize findings, (d) external, intervening events create bias, (e) uncontrollability is exacerbated if there is too much discretion and community choice across sites, and (f) you cannot clearly attribute effects to interventions because of these 'interferences'. One of the 'interferences' is the actual participation and partnership which is acknowledged as having an effect on the interventions and outcomes of the research (Springett & Wallerstein, 2008).

It is clear that *validity* requirements are an ongoing challenge, given the 'messiness' of community-engaged research. Although this remains an area of contestation in mainstream research circles, particularly within medical arena, many researchers have found ways to address the concerns usually raised (e.g. Bradbury & Reason, 2008; Mertens, 2005; Torres-Harding et al., 2004). These strategies are described below.

In addition to the various methodological challenges identified above, community research has *political* inherency (Mertens, 2005) and is *messy*, requiring longer time-frames and high levels of flexibility. Historical abuse of research, and political agendas and dynamics in the community, can also hamper the research (Farquhar & Wing, 2008).

Springett and Wallerstein (2008) raise further specific challenges:

- Pre-existing goals of the research can constrain issue selection.
- It often involves a great deal of time and resources which need to be built into research protocols and funding grants.
- Funders hold back on funding unpredictable processes and outcomes.
- Researchers often lack all the skills required, and the community does not always think 'critically'.
- The larger the reach of a project, the more difficult it is to ensure a democratic process.
- It is difficult to control the 'coming and going' of people in the project.
- The relationship between researcher and researched, the balance between expert and lay involvement, and insider/outsider dynamics are all challenges.

A further area of concern for many in public health and other circles is that of *ongoing colonisation* through research and knowledge production. Foucault's theory of "governmentality" (Foucault, 1980) describes the powerful conditions that influence individuals and populations to actively self-regulate their own behavior in alignment with scientific and evidence based prescription of health. This insight unveils the hidden mechanism of power in standard approaches to community engagement. Through the privileging of evidence-based interventions, health becomes an "*ethical imperative*", requiring individuals to regulate their behavior and reshape their selves in keeping with new bio-medical and public health knowledge. Those who acquire these science sanctioned behaviours earn the status of sanitary citizens (Briggs & Mantini-Briggs 2003), individuals deemed to possess modern medical understandings of the body, health, and illness, practice hygiene, and depend on doctors and nurses when they are sick (Briggs, 2001; Ong, 1987). People who are assumed to be incapable of accepting this modern medicalised relationship to the body, hygiene, illness, and healing—or who reject this subjectivity—become unsanitary subjects (Briggs, 2005). These terms incorporate what have been referred to as biomedical citizens (Ong 1995, Shah 2001) and, at the same time, draw attention to the broader moral, social, political, and cultural meanings that shape how social responsibility is defined in terms of health. These discursive mechanisms are a return to the historical uses of medicine and public health in colonisation.

The *literature* consulted revealed that researchers working within this paradigm have found a variety of *strategies to address the challenges* highlighted above (Bradbury & Reason, 2008; Farquhar & Wing, 2008; Flicker et al., 2008; Mertens, 2005; Minkler & Baden, 2008; Springett & Wallerstein, 2008;). This includes*:*

- Ensuring that the community's needs and agenda's are addressed
- Developing and utilising participatory leadership skills
- Facilitating optimal participation during each step of the research process
- Balancing the various tensions in this type of research
- Using rigorous tools and techniques, using locally and culturally appropriate methods and instruments
- Using multiple methods
- Developing different kinds of validity measures that cover the requirement for rigour and quality
- Working within a strict code of conduct for ethics purposes

- Translating the research for both dissemination and utilization purposes, in languages and formats relevant to the people concerned

Springett and Wallerstein (2008) stress the need to *balance* the requirements of participation, the use of rigorous tools and techniques, and the practical demands of the real world, while retaining the values of social action. They argue that this requires good participatory *leadership* skills to hold the project.

As mentioned earlier, a number of researchers have found ways to address the concerns usually raised around *validity* (e.g. Bradbury & Reason, 2008; Mertens, 2005; Torres-Harding et al., 2004). This includes emphasizing that there are different kinds of validity (e.g. methodological validity, interpersonal validity, consequential validity, and multicultural validity) which need to be taken into account. Bradbury and Reason (2008) argue for the alternative use of the term *quality* which is linked to relational praxis, reflexive or practical outcome, plurality of knowing, conceptual and theoretical integrity, extending ways of knowing, methodological appropriateness, and engaging in significant work.

With regard to addressing the *participation* challenges in each of the research steps, Minkler and Baden (2008) make the following suggestions:

- Research question selection: The question needs to meet the needs from both sides
- Instrument and research design: Community input is necessary and one often needs to forego an emphasis on 'control'
- Ethical review and informed consent: This needs to fit the local culture
- Data collection: Methods need to be developed with the community and adjusted where needed
- Data analysis and interpretation: Involvement of community in analysis is not always possible or appropriate, but it is optimal and can be very successful
- Dissemination and use of findings: Various methods, including public approaches, need to be pursued for dissemination, and the focus on 'action' can create tensions between academics and community because researchers generally do not get involved in the 'action'.

Flicker et al. (2008) provide a very useful framework for *ethical considerations* in CBPR research, providing guidelines for CBPR ethics boards, based on an analysis across 30 schools of public health in the USA. They argue that, in addition to normal ethics guidelines (e.g. autonomy, nonmaleficence, beneficence, justice), the principles of CBPR must also be made visible. This includes ensuring that there is a clear Terms of Reference or Memorandum of Understanding between the research partners. They also argue that the process of the research needs to be documented.

5. Reflections on CBPR challenges in practice

The following *challenges have been found to be particularly relevant to the authors* based on their varied experiences in marginalised and historically oppressed communities. Many of these challenges can be located within researcher-community relationship dynamics, particularly in relation to participation and community consent, knowledge and power, resources and privilege, general community dynamics, and specific research considerations.

5.1 Knowledge and power

Challenges relating to knowledge and power include the following key issues:

- Researchers and communities usually have different interests in knowledge production, with the latter usually being more interested in solving particular practical problems.
- Academic researchers are usually viewed as being the experts who have 'scientific knowledge', and this creates imbalances in power relations. This perceived academic expertise may silence others' voices.
- These imbalances are supported by government and other powerful structures which favour academic language and a particular form of scientific knowledge that reflects an Anglo-Euro-American worldview.
- There is generally a lack of acknowledgement and honouring of all kinds of knowledge systems.

5.2 Resources and power

Challenges relating to resources, which are linked to access to power in society, include the following:

- There is clearly inequitable access to all resources relating to research, with local communities having minimal, if any, access to the required financial and skill-resources.
- Academic researchers usually have access to the resources for the research, and therefore the balance of power.
- The role and interests of funders play a key role, often overtly or covertly guiding the community research agenda.

5.3 Participation and power

Who is involved in the research, and how, is a key consideration when conducting CBPR. Some of the key challenges relating to participation are listed below.

- On a practical level, the question of *who initiates* the project is an important one as it often reflects and perpetuates imbalanced power relations in the research partnership.
- A related question is 'whose *research agenda'* is on the table. A challenge for community-based researchers is accepting that everything that has statistical significance may not be relevant to community outcomes. Community based researchers are often confronted with meaningful findings from mixed method studies that could advance the communities objectives but not their respective fields. Similarly researchers are challenged with the tensions between community capacity (e.g. empowerment) and professional outcomes (i.e. publication).
- The question of *who is involved* in the research process is also an important question. This includes challenges relating to identifying the 'right' people, and then including them fully in the process.
- Once one has all the relevant people 'around the table', a central challenge relates to *facilitating optimal participation*. This includes mediating ongoing power dynamics and tensions. It is important to note that the levels of participation and control are never

static, and expectations of levels of participation vary by levels of ownership. Furthermore, power and control shifts with function and content.

5.4 Community dynamics

Challenges relating to community dynamics are always present. This includes the following issues:

- Conducting research in and with the community always involves formal and informal political dynamics which can interfere with the research process.
- This includes challenges relating to cultural diversity which is a reality in any modern community.
- It is necessary to understand the current power dynamics, including a historical and contemporary analysis of positions and other forms of power.
- The challenge of understanding one's own position of power in these dynamics is crucial. This relates to the challenge of reflexivity – self-reflection on one's position within power relations.
- The community's possible history of oppression (a reality in all the contexts of practice of the authors of this paper) needs to be understood and healed and utilised for the purposes of both personal and political empowerment. The collective memory of oppression is often transmitted inter-generationally through folklore, fables, or direct instruction and serves as the foundation for cultural mistrust that influences health behaviour. Researchers rarely take the time to hear the stories of marginalisation, exploitation, and often terror that serve as determinants to a community's health behaviours. Balancing historical realities with contemporary outcomes often challenges researchers who exist in a "pseudo objective" scientific bubble. Communities approach researchers with an ascribed set of characteristics based on previous exposure to exploitation.
- There are many challenges relating to community empowerment. This includes both fostering personal and collective agency, and managing an understandable but often destructive reaction of entitlement.
- Sustainability relating to community development is also a key challenge. Many of the above mentioned challenges relate to this issue.

5.5 Research methodology challenges

Challenges relating to the actual research methodology employed in a CBPR project are real, and need to be addressed.

- Challenges relating to *control* have been well articulated in the literature briefly discussed above. The tensions of control have to be managed in an ongoing way.
- Acknowledgement of the *messiness* of this approach to research is also important. This includes being aware of the need to be flexible in the process of the research.
- Challenges relating to *research translation* include ensuring that the action envisaged actually happens, and is sustained in some way within the community concerned.
- The urgency that communities face to find solutions to challenges usually do not fit within the *timeframe* of the research process. Researchers often enjoy the privilege of

exploration because their wellbeing is often not intrinsically dependent on the answer to a research question. Time to think deeply and employ systematic methods to answering questions is a privilege that researches take for granted. The challenge is for researchers to respect the cultural and pragmatic function of time when participating in research with communities.

The *suggestions of ways of addressing these challenges* identified in the literature (refer the previous section) are relevant in the varied contexts of practice of the authors of this paper. In addition to the points raised by others, the following important considerations and recommendations are offered, based on reflections of our own practices in historically oppressed communities that are still marginalised in one way or another. It should be noted that most of these recommendations relate directly to the *challenges of power* highlighted in the previous section. Given the historical (and contemporary) realities of these communities, this is not surprising.

- Given the power dynamics relating to knowledge generation, researchers should engage community stakeholders in a *dialogue* that genuinely honours the different forms of knowledge 'around the table'.
- This includes *decolonising knowledge* systems, and honouring and providing opportunities to engage with different knowledge systems and research approaches. In meeting these challenges it is important to *honour both academic and community-bedded or indigenous knowledges* in order to realign the power dynamics.
- With regard to the privileging of academic knowledge, it is important to create spaces for *postcolonial and hybrid knowledge production*, and including culturally supported interventions, indigenous theories, and decolonising methodologies.
- With regard to incompatible discourses between academia and community, we need to broaden the discourse to include *'life world' cultural and social meanings.*
- We need to shift power between universities and communities through *bidirectional* learning, shared resources, collective decision making and outcomes beneficial to the community, including the co-discovery and promotion of community scholarship.
- A researcher's ability to *acknowledge privilege and share power* in community settings is a powerful tool in CBPR. Effective engagement with communities requires the shedding of the formal academic training which over emphasizes individual contributions in favour of a more ecological perspective of community health. Shifting paradigms to more collective strength-based approaches is probably one of the most important methods of addressing the challenges to CBPR.
- Our collective experience has highlighted the important of *reflexivity*. This self-reflective process is discussed in more detail in a forthcoming publication (Duran, Lazarus, Caldwell & Bulbulia). For the purpose of this paper, it is important to emphasise that we need to start with ourselves, examining our own interests and position in the research process and community dynamics. This includes constantly reflecting on 'whose research agenda' is being promoted and pursued, ensuring that community interests are always at the forefront.
- *Acknowledgement of cultural diversity* is central to working with the inevitable differences in any community. Dealing with these differences includes advocating for inclusive, accountable and transparent processes, and promoting ongoing facilitated dialogue which, at times, may require various conflict management strategies.

- *Facilitating empowerment* requires employing strategies to foster both personal and collective agency, within the context of collective accountability.
- *Engagement* requires addressing various *gaps*. This includes general gaps in accessing resources, including information and knowledge. It is therefore important to redress current inequalities relating to accessing information and knowledge, as well as access to inclusive processes in knowledge construction.
- *Sustainability* relating to community development is a challenge which has to be addressed. It is important to build this capacity during project development. All CBPR project should be guided by general community development principles which have been developed over time, in our own and other similar contexts.
- To ensure that *research translation* does occur, that is, that the action proposed is followed through, it is important that appropriate forms of research dissemination are developed. This includes ensuring that knowledge dissemination and production is mutually owned and respected. Proper 'follow through' also means that resources must be allocated to 'action' phases of the research intervention. This means that researchers have to consider their commitment to the community(ies) concerned beyond normal academic requirements!

A central challenge in CBPR is to conduct rigorous research that is both culturally responsive and scientifically sound. Putting these two aspects together requires creative and innovative strategies, and fundamentally, a deep respect for diverse ways of seeing and doing, and embracing different knowledge systems. Research needs to be demystified and decolonized and made accessible, appropriate and relevant to the community context.

6. In conclusion

> *No matter how ignorant a person is, there is one thing he / she knows better than anybody else, and that is where the shoes pinch his / her own feet, and because it is the individual that knows his / her own troubles even if he / she is not literate or sophisticated in other respects, every individual must be consulted in such a way, actively, not passively, that he himself / herself, becomes a part of the process of authority, of the process of social control, that his / her needs and wants have a chance to be registered in a way that they count in determining social policy*
> (John Dewey)

Central to the arguments in this discussion on CBPR as a research and action strategy within public health is the need for *respect*. Engaging with this value and principle, within ourselves and with our research partners, provides a firm basis for the development of related values and principles which should guide our research practice. Reflection on ourselves and our practice, emphasised in this chapter, must include an honest engagement with our position in the various dynamics of power within the research relationship. We are all challenged to recognise that our individual and collective development and liberation is a joint journey which requires us to walk and talk together in mutual respect and wonder.

> *If you have come to help me, you are wasting your time, but if you have come because your liberation is bound to mine, then let us work together*
> (Lila Watson: Aboriginal woman leader)

7. References

Attree, P. & French, B. (2007). *Testing theories of change associated with community engagement in health improvement and health inequalities reduction.* Report prepared for NICE (Available on request from p.attree@lancaster.ac.uk).

Bradbury, H. & Reason, P. (2008). Issues and choice points for improving the quality of action research. In Minkler, M. & Wallerstein, N. *Community-based participatory research for health: From process to outcomes*. San Francisco, Jossey-Bass.

Briggs, C. L. (2001). Modernity, Cultural Reasoning, and the Institutionalization of Social Inequality: Racializing. *Comparative Studies in Society & History, 43*(4), 665.

Briggs, C. L. (2005). Communicability, Racial Discourse, and Disease. *Annual Review of Anthropology, 34*(1), 269-291.

Cornwall, A. & Jewkes, R. (1995). What is participatory research? *Soc Sci Med, 41*(12), 1667-1676.

Durlak et al (2004). In Jason, L.A., Keys, C.B., Suarez-Balcazar, Y.S., Taylor, R.R. & Davis, M.I. (2004). *Participatory community research: Theories and methods in action*. Washington: APA.

Fals-Borda, O., & Rahman, M. A. (1991). *Action and knowledge: Breaking the monopoly with participatory action research*. New York/London: Apex Press Intermediate Technology Publications.

Farquhar, S.A. & Wing, S. (2008). Methodological and ethical considerations in community-driven environmental justice research. In Minkler, M. & Wallerstein, N. *Community-based participatory research for health: From process to outcomes*. San Francisco: Jossey-Bass.

Felner, D., Felner, T.Y. & Silverman, M.M. (2000). Prevention in mental health and social intervention: Conceptual and methodological issues in the evolution of science and practice of prevention. In J. Rappaport & E. Seidman (Eds). *Handbook of Community Psychology*. New York: Kluwer Academic/Plenum Publishers.

Flicker, S., Travers, R., Guta, A., McDonald, S. & Meagher, A. (2008). Ethical review of community-based participatory research: Considerations for Institutional Review Boards. In Minkler, M. & Wallerstein, N. (Appendix G). *Community-based participatory research for health: From process to outcomes*. San Francisco, Jossey-Bass.

Foucault, M. (1980). *The history of sexuality, Volume 1: An introduction* (R. Hurley, Trans.). New York: Vintage/Random House.

Fried, C.S., Reppucci, N.D. & Woolard, J.L. (2000). Violence prevention. In J. Rappaport & E. Seidman (Eds). *Handbook of Community Psychology*. New York: Kluwer Academic/Plenum Publishers.

Green, L. A., & Hickner, J. (2006). A short history of primary care practice-based research networks: from concept to essential research laboratories. *J Am Board Fam Med, 19*(1), 1-10.

Hashagen, S. (2002). *Models of community engagement.*

Hawe, P., Shiell, A. & Riley, T. (2009). Theorising interventions as events in systems. *American journal of community psychology, 43*(3-4), 267-276.

Heller, K., Price, R.H., Reinharz, S., Riger, S., Wandersman, A. & D'Aunno, T.A. (1984). *Psychology and community change*. Illinois: The Dorsey Press.

Herbertson, K., Ballesteros, A.R., Goodland, R. & Munilla, I. (2009). *Breaking ground: Engaging communities in extractive and infrastructure projects.* World Resources Institute, Washington, DC.

Israel, B. A. (2005). *Methods in community-based participatory research for health* (1st ed.). San Francisco, CA: Jossey-Bass.

Israel, B. A., Schulz, A. J., Parker, E. A., & Becker, A. B. (2001). Community-based participatory research: policy recommendations for promoting a partnership approach in health research. *Educ Health (Abingdon), 14*(2), 182-197.

Jason, L.A., Keys, C.B., Suarez-Balcazar, Y.S., Taylor, R.R. & Davis, M.I. (2004). *Participatory community research: Theories and methods in action.* Washington: APA.

Kramer, & Specht, (1983, 3rd Edition). *Readings in Community Organization Practice*. New Jersey: Prentice-Hall.

Lewin, K. (1938). *The conceptual representation and the measurement of psychological forces*. Durham, N. C: Duke university press.

Lewin, K. (1951). *Field theory in social science; selected theoretical papers* ([1st ed.). New York: Harper.

Lewin, K. (1997). *Resolving social conflicts ; Field theory in social science*. Washington, DC: American Psychological Association.

Lewin, K., & University of Michigan. Research Center for Group Dynamics. (1975). *Field theory in social science: Selected theoretical papers*. Westport, Conn.: Greenwood Press.

Mertens, D.M. (2005, 2nd Edition). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods*. California: Sage Publications.

Minkler, M. & Baden, A.C. (2008). Impacts of CBPR on academic researchers, research quality and methodology, and power relations. In Minkler, M. & Wallerstein, N. *Community-based participatory research for health: From process to outcomes*. San Francisco, Jossey-Bass.

Minkler, M. & Hancock, T. (2008). Community-driven asset identification and issue selection. In Minkler, M. & Wallerstein, N. *Community-based participatory research for health: From process to outcomes.* San Francisco, Jossey-Bass.

Minkler, M. & Wallerstein, N. (2008, 2nd Edition). *Community-based participatory research for health: From process to outcomes.* San Francisco: Jossey-Bass.

NACCHO. (2001). *Fourteen policy principles for advancing collaboration among and between tribal communities and surrounding jurisdictions*. Retrieved October 25,2011,from http://www.naccho.org/topics/infrastructure/turningpoint/upload/TP-policy-principles.pdf

Ong, A. (1987). *Spirits of Resistance and Capitalist Discipline*. New York: State University of New York Press.

Pokorny et al (2004). In Jason, L.A., Keys, C.B., Suarez-Balcazar, Y.S., Taylor, R.R. & Davis, M.I. (2004). *Participatory community research: Theories and methods in action.* Washington: APA.

Popay, J. (2006). *Community engagement and community development and health improvement: A background paper for NICE* (available on request from antony.morgan@nice.org.uk).

Rappaport, J. & Seidman, E. (2000). *Handbook of community psychology.* New York: Kluwer Academic/Plenum Publishers.

Reason, P., & Bradbury, H. (2001). *Handbook of action research : participative inquiry and practice*. London, Thousand Oaks, Calif.: SAGE.

Reason, P., & Bradbury, H. (2008). *The Sage handbook of action research : participative inquiry and practice* (2nd ed.). London, Thousand Oaks, Calif.: SAGE Publications.

Revenson, T.A. & Schiaffino, K.M. (2000). Community-based health interventions. In J. Rappaport & E. Seidman (Eds). *Handbook of Community Psychology*. New York: Kluwer Academic/Plenum Publishers.

Rifkin, S.B., Lewando-Hundt, G. & Draper, A.K. (2000). *Participatory approaches in health promotion and planning: A literature review*. London: Health Development Agency.

Roesch, R. & Carr, G. (2000). Psychology in the international community: Perspectives on peace and development. In J. Rappaport & E. Seidman (Eds). *Handbook of Community Psychology*. New York: Kluwer Academic/Plenum Publishers.

Rossi, P.H., Lipsey, M.W. & Freeman, H.E. (2004, 7th Edition). *Evaluation: A systematic approach*. California: Sage Publications.

Schulz, A. J., Parker, E. A., Israel, B. A., Allen, A., Decarlo, M. & Lockett, M. (2002). Addressing social determinants of health through community-based participatory research: the East Side Village Health Worker Partnership. *Health Educ Behav, 29*(3), 326-341.

Springett, J. & Wallerstein, N. (2008). Issues in participatory evaluation. In Minkler, M. & Wallerstein, N. *Community-based participatory research for health: From process to outcomes*. San Francisco, Jossey-Bass.

Stoeker, R. (2005). *Research methods for community change: A project-based approach*. California: Sage Publications.

Taylor et al (2004). In Jason, L.A., Keys, C.B., Suarez-Balcazar, Y.S., Taylor, R.R. & Davis, M.I. (2004). *Participatory community research: Theories and methods in action*. Washington: APA.

Torres-Harding et al (2004). In Jason, L.A., Keys, C.B., Suarez-Balcazar, Y.S., Taylor, R.R. & Davis, M.I. (2004). *Participatory community research: Theories and methods in action.* Washington: APA.

Trickett, E. J. Community-based participatory research as worldview or instrumental strategy: Is it lost in translation(al) research? *Am J Public Health, 101*(8), 1353-1355.

Wallerstein, N., & Duran, B. (2008). The theoretical, historical, and practice roots of CBPR. In M. Minkler & N. Wallerstein (Eds.), *Community Based Participatory Research for Health* (2nd ed). San Francisco: John Wiley and Sons.

Wallerstein, N., Oetzel, J., Duran, B., Tafoya, G., Belone, L. & Rae, R. (2008). What predicts outcomes in CBPR? In Minkler, M. & Wallerstein, N. *Community-based participatory research for health: From process to outcomes*. San Francisco, Jossey-Bass.

Westfall, J. M., Mold, J., & Fagnan, L. (2007). Practice-based research: "Blue Highways" on the NIH roadmap. *JAMA, 297*(4), 403-406.

16

Nature Therapy and Preventive Medicine

Juyoung Lee[1], Qing Li[2], Liisa Tyrväinen[3], Yuko Tsunetsugu[4],
Bum-Jin Park[5], Takahide Kagawa[4]and Yoshifumi Miyazaki[1]
[1]Center for Environment, Health and Field Sciences, Chiba University,
[2]Nippon Medical School,
[3]Finnish Forest Research Institute(METLA),
[4]Forestry and Forest Products Research Institute,
[5]Chungnam National University,
[1,2,4]Japan
[3]Finland
[5]Korea

1. Introduction

Five million years passed before humans evolved into what we are today. Therefore, more than 99.99% of our evolutionary history was spent in natural environments, assuming that urbanization can be defined as a postindustrial revolution development. We have become the species we are today, living in a modern civilization, through a process of evolution within a natural environment. Human bodies are made so as to adapt to nature. However, terms that we hear today, such as "technostress (Brod, 1984)," indicate that artificialization, which is the process of society and our lives being intruded by technologies of different types, is occurring so rapidly that we now experience stressful situations and are forced to deal with the resultant pressures.

If, under such circumstances, we receive a nature-based stimulus through nature therapy, we may become aware of what we really are. We may have the chance to relax and be very comfortable. This is accomplished without logical thought. Instead, we intuitively perceive the world through the five senses. Because this process cannot be described in words, physiological indicators play an important role.

Although we are now living in a society characterized by urbanization and artificialization, our physiological functions are still adapted to to nature (Miyazaki et al., 2011). Because of this discrepancy between our body requirements and our manner of living, our stress levels are always very high and our sympathetic nervous system is excessively stimulated. In many cases, lowering elevated stress levels to a point where the body can function properly is an immediate necessity. This is exactly where nature therapy comes in. As the body approaches the "expected" natural state of well-being, immune functions are enhanced and disease resistance improves. In modern medicine, for example, we treat ourselves with antibiotics when we contract pneumonia, expecting them to relieve symptoms. In contrast, nature therapy causes a "nonspecific effect" whereby our bodies become resistant to pneumonia as a result of increased immune function induced by relief of body stress.

Our recent field research investigating immune responses to natural environments demonstrated that immune functions are enhanced by contact with forest environments. Middle-aged employees who had no subjective symptoms but who complained of lack of energy and decreased immune functionvolunteered to participate in our study. We examined natural killer cell activity (NK activity) as an indicator of immune function, particularly as an indicator of anticancer activity. After the second day walk in the local forest, NK activity was enhanced by 56% in these subjects, and normal immune functions were restored. A statistically significant increase of 23% was maintained for 1 month even after these volunteers had returned to urban life, clearly illustrating the preventive medical effect of nature therapy (Li et al., 2007a, 2008a, 2008b, 2010).

A great deal of attention is now being paid to nature therapy to scientifically identify synchronizations between humans and nature using data on the physiological effects of relaxation. In this chapter, based on human physiological data, we outline the present state of research on nature therapy and its scientific evidence from the viewpoint of preventive medicine. In addition, individual differences in the physiological effects of nature therapy and current initiatives on the relationship between nature and human health are discussed.

2. Nature therapy and well-being

2.1 Health and well-being

People have a growing interest in ways to improve their own health, which perhaps reflects the stresses of modern society. How can we achieve a sense of well-being? Obviously, being healthy does not simply mean freedom from disease.

Various definitions of well-being can be found nowadays; one understanding of the word "healthy" is "the state in which an individual fully displays the abilities he or she has or is born with." Because it differs from one individual to another, a healthy state can be maintained even when living with a physical disability. Thus, well-being is relative rather than absolute. Even more important is the conviction that being healthy is not just an ideal "picture" in itself but rather a process through which a positive and constructive life can be led. In other words, health should be a "means" not a "goal."

2.2 Nature and comfort

"Comfort" is a common term in daily life, yet no set definition is available in the academic community. One understanding of comfort is the "harmonization of rhythm between human beings and the environment." During everyday life, a feeling of comfort can be experienced if our rhythms are synchronized with those of the environment. This comfort is sometimes reflected in phrases such as "We hit it off immediately," "We get on well with each other," or "I like the atmosphere." For example, when I feel that the audience is listening to my lecture and that our intentions are synchronized, I feel comfortable with the idea of talking some more. Humans can achieve peace in natural surroundings, gaining comfort as a result of synchronization between us and Mother Nature.

Masao Inui categorized comfort using the terms "passive comfort" and "active comfort." "Passive comfort" is rooted in the desire for safety and the elimination of discomfort. When assessing this category of comfort, it should not be difficult to reach consensus because individual preferences are not involved. In contrast, “active comfort” is rooted in the desire for personal growth and the urge to achieve something extra. Of course, personal aims can change and it can be difficult to reach consensus even within oneself. Active comfort is needed most in today’s society. Although passive comfort is the basic minimum that must be achieved, in future research on comfort, active comfort will be the main subject of study as its assessment is already attracting the interest of many researchers.

2.3 Forest therapy

Based on these effects of nature on human well-being, Japan’s Forestry Agency startedand provided the blueprint for the “forest therapy stations” project in 2005. The term “forest therapy” was coined during the process of developing this project, and the underlying idea was borrowed from the term “aromatherapy.”

The blueprint for forest therapy was prepared as a result of rapid advances made over the last 3 years in assessment of physiological techniques to achieve relaxation (Lee et al., 2011a; Tsunetsugu et al., 2010). From 2005–2011, the team organized for the study [primarily from the Center for Environment, Health and Field Sciences, Chiba University and the Forestry and Forest Products Research Institute (FFPRI)] performed experiments over a period of about 1 week in 48 different forests at various locations throughout Japan, ranging from the large northern island of Hokkaido to Okinawa in the south. Measured variables (endpoints) included stress hormone (cortisol) levels in saliva and autonomic nervous activity (sympathetic and parasympathetic) monitored by heart rate fluctuation, blood pressure, and heart rate. In addition, for the first time in the world, a method was developed to monitor prefrontal cortex activity in the brain using near-infrared spectroscopy in the field. Measurements of forest phytoncides, urban exhaust fumes, temperature and humidity, illuminance, wind velocity, and negative (minus) and positive (plus) ions were also performed. By conducting these studies with human volunteers at forest sites throughout Japan, we were able to confirm the physiological effects of relaxation and provide scientific evidence for the benefits of forest therapy.

3. Methodology for evaluating health-related effects

3.1 Physiological relaxation

To investigate the physiological effects of actual natural environments, all measurement and sampling was performed in the field. Physiological data were collected through measurements of salivary cortisol concentration, blood pressure (systolic and diastolic), pulse rate, and heart rate variability.

Salivary cortisol, a stress hormone, was analyzed to determine stress responses to natural environments. Cortisol is released by the hypothalamic–pituitary–adrenal (HPA) axis in response to stress (Seplaki et al., 2004) and is a reliable indicator of endocrine stress responses (Kirschbaum & Hellhammer, 1989). Its release in response to stress is immediate and is highly associated with the free cortisol fraction in blood (Kirschbaum & Hellhammer, 1994). The sampling procedure is very simple and does not affect the cortisol values (Kirschbaum & Hellhammer, 1989). In response to a stressor, the excretion of cortisol (Pruessner et al., 1999), blood pressure, and pulse rate generally increases (Sluiter et al., 2000). In this study, saliva samples were obtained using a salivette device (No. 51.1534; Sarstedt, Numbrecht, Germany), immediately frozen, and transported to the laboratory (SRL, Inc., Tsukuba, Japan) for analysis of cortisol concentrations.

Pulse rate and blood pressure (systolic and diastolic blood pressures) as indices of autonomic nervous system activity were measured by the oscillometric method using a digital blood pressure monitor (HEM1000; Omron, Japan). Heart rate variability (HRV), an indicator of human autonomic activity (Kobayashi et al., 1999), was measured using a portable electrocardiograph (Activtracer AC-301A, GMS, Tokyo, Japan). HRV data were obtained at various frequencies using an HRV software tool (MemCalc/win, GMS). For real-time analysis of HRV, interbeat (R–R) intervals were obtained in 1-min segments using the maximum entropy method. Variance of the two major spectral components of HRV was calculated: the low-frequency (LF; 0.04–0.15 Hz) band and the high-frequency (HF; 0.15–0.4 Hz) band (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996). The LF/HF ratio in R–R interval variability was also assessed. We used the HF data as an index of parasympathetic nervous activity (Cacioppo et al., 1994) and the LF/HF ratio as an index of sympathetic nervous activity (Weise & Heydenreich, 1989).

3.2 Immune function

Effects of nature therapy on human immune function were evaluated by measurement of NK activity, the number of NK cells, and intracellular levels of anticancer proteins, including perforin, granulysin (GRN), granzyme A (GrA), and GrB in peripheral blood (Li et al., 2007a, 2008a, 2008b, 2010; Li, 2010a; Li & Kawada, 2011).

3.2.1 NK activity

Peripheral blood lymphocytes (PBLs) were separated from peripheral blood using a BD Vacutainer CPT tube, and adjusted to 4 × 10^6 cells/ml. NK activity was assayed according to a standard method (Li et al., 2007a, 2008a, 2008b, 2010; Li, 2010b). In brief, K-562 target cells were labeled with a sodium ^{51}Cr-chromate solution for 60 min at 37°C in 5% CO_2 and washed 4 times in RPMI-1640 containing 10% fetal bovine serum (FBS). The target cells were placed in round-bottomed 96-well microplates. Effector cells (PBLs) at 4 × 10^6, 2 × 10^6, and 1 × 10^6 cells/ml in 100 μl were then added to the wells in triplicate at E:T ratios of 40:1, 20:1, and 10:1, respectively. Following a 4-h incubation period at 37°C in 5% CO_2, the microplates were centrifuged and 100 μl of supernatant from each well was collected and measured in a gamma counter. NK activity was then calculated as described previously (Li et al., 2007a, 2008a, 2008b, 2010; Li, 2010b).

3.2.2 NK cells and perforin-, GRN- and GrA/B-expressing lymphocytes by flow cytometry

NK cells in PBLs were stained with phycoerythrin (PE)–CD16 monoclonal antibody and PE–mouse IgG1 as a negative control for 30 min in the dark. The cells were then fixed/permeabilized with Cytofix/Cytoperm solution for 20 min at 4°C. Intracellular perforin and GrA/B were stained with fluorescein isothiocynate (FITC)–antihuman perforin and FITC–GrA/B antibodies. As negative controls, FITC–IgG2b was used for perforin and

FITC–IgG1 for GrA/B for 30 min at 4°C according to the manufacturer's instructions. Intracellular GRN was stained with a rabbit anti-human GRN polyclonal antibody using rabbit serum as the negative control after fixation/permeabilization with Cytofix/Cytoperm solution, and then stained with FITC–goat antirabbit IgG for 30 min at 4°C in the dark (Li et al., 2007a, 2007b, 2008a, 2008b, 2010; Li, 2010b). After staining, the cells were washed twice with fixative solution and once with PBS containing 1% FBS. Flow cytometric analysis was performed with a FACScan flow cytometer as described previously (Li et al., 2007a, 2007b, 2008a, 2008b, 2010; Li, 2010b).

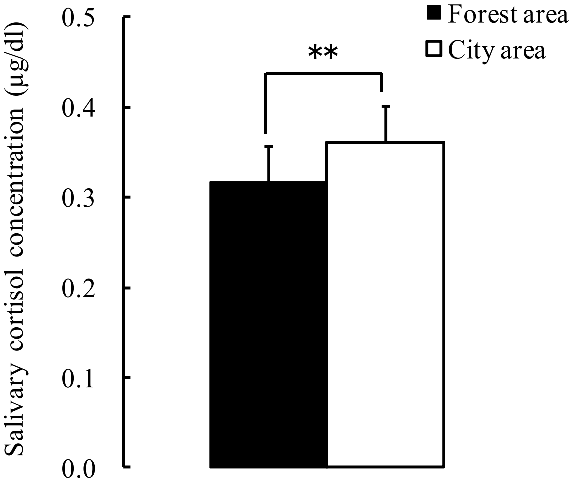
4. Evidence-based approach to health benefits of natural environments

4.1 Field experiments

4.1.1 Nature and physiological relaxation

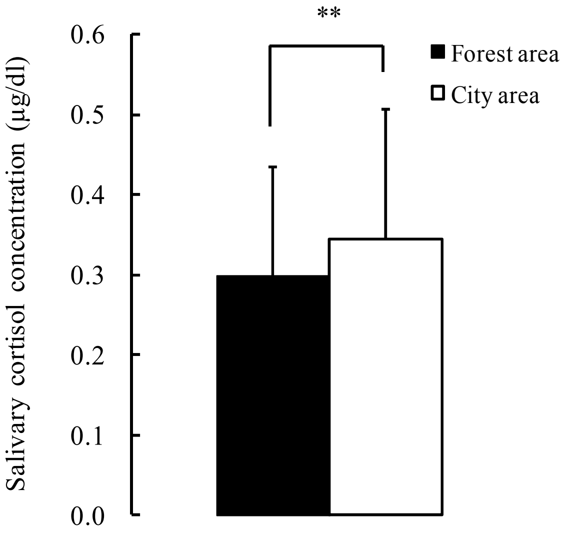
Investigation of human physiological responses in actual field sites is extremely valuable because it can reveal the total effects of the surrounding environment (Lee et al., 2009, 2011b, 2011c; Matsunaga et al., 2011; Park et al., 2007, 2008a, 2008b, 2009; Tsunetsugu et al. 2007b). A field study provides more important information regarding the effects of real environments than an indoor study. From 2005, scientific field data on physiological responses to natural environments have been accumulated and compared with responses in artificial city environments. Here, we present some of this physiological data from experiments conducted in Japanese forest sites. Subjects were assessed after viewing and walking in the forest area.

The concentration of salivary cortisol was significantly decreased when subjects were in the forest area (12.4% decrease after viewing; 15.8% decrease after walking) compared with when they were in the city area (Figs. 1 and 2, respectively).



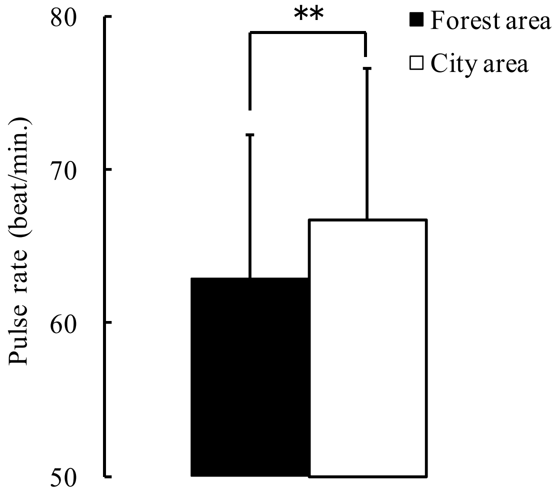
N = 385; Mean ± SD; ** $p < 0.01$; p-value by t-test. (Source: Park et al., 2011)

Fig. 1. Change in salivary cortisol concentration after forest viewing.



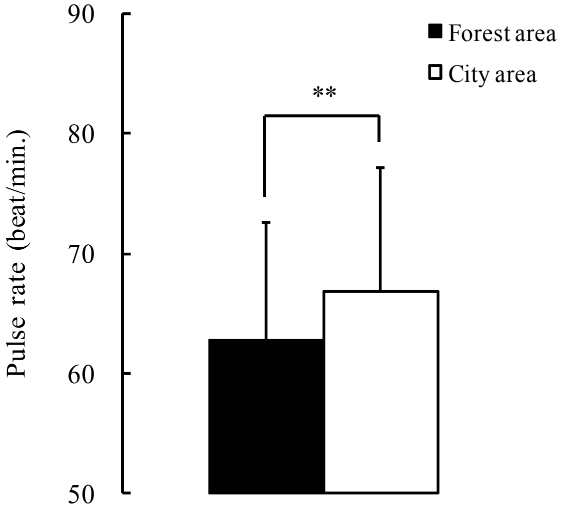
N = 74; Mean ± SD; ** $p < 0.01$; p-value by t-test. (Source: Park et al., 2010)

Fig. 2. Change in salivary cortisol concentration after forest walking.



N = 397; Mean ± SD; ** $p < 0.01$; p-value by t-test. (Source: Park et al., 2011)

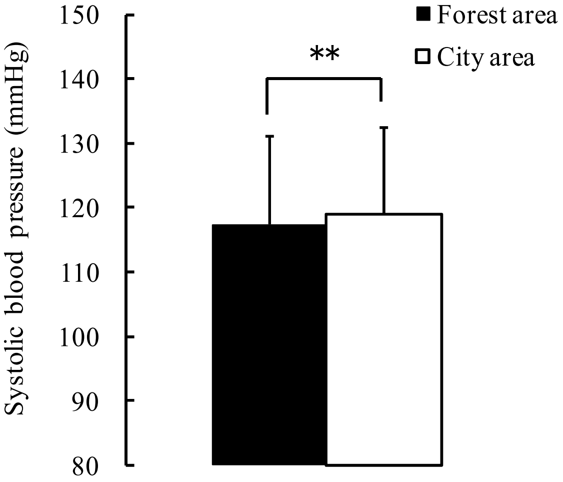
Fig. 3. Change in pulse rate after forest viewing.



N = 75; Mean ± SD; ** p < 0.01; p-value by t-test. (Source: Park et al., 2011)

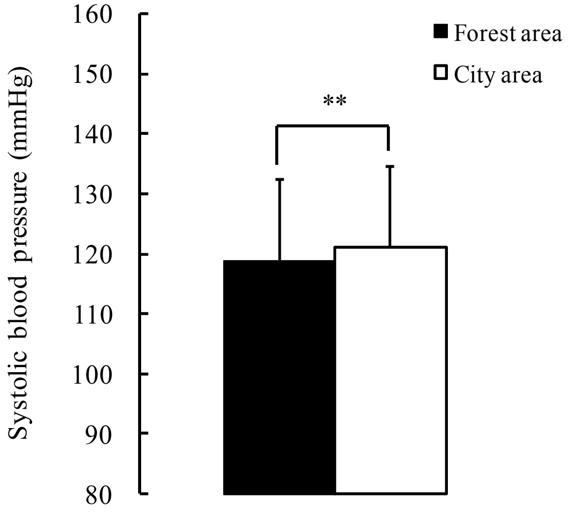
Fig. 4. Change in pulse rate after forest walking.

Figures 3 and 4 show that the average pulse rate was significantly reduced when subjects were in forest environments compared to that when they were in city environments (5.8% decrease after viewing; 3.9% decrease after walking). Similar characteristics in response to the two different modes of environmental stimulation were observed in systolic blood pressure. The average systolic blood pressure was significantly lower in the forest environment than in the city environment (1.4% decrease after viewing; 1.9% decrease after walking; Figs. 5 and 6).



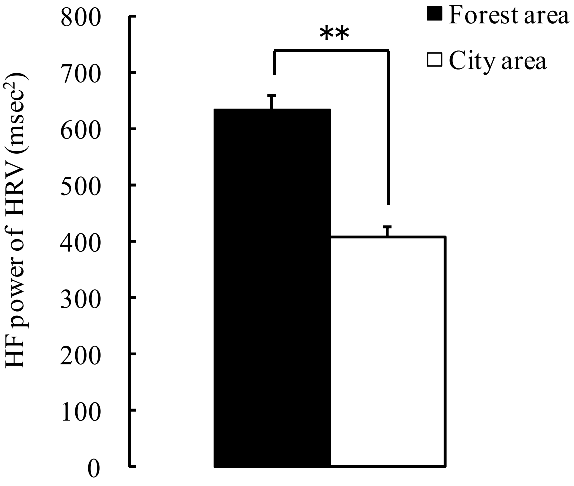
N = 397; Mean ± SD; ** p < 0.01; p-value by t-test. (Source: Park et al., in press)

Fig. 5. Change in systolic blood pressure after forest viewing.



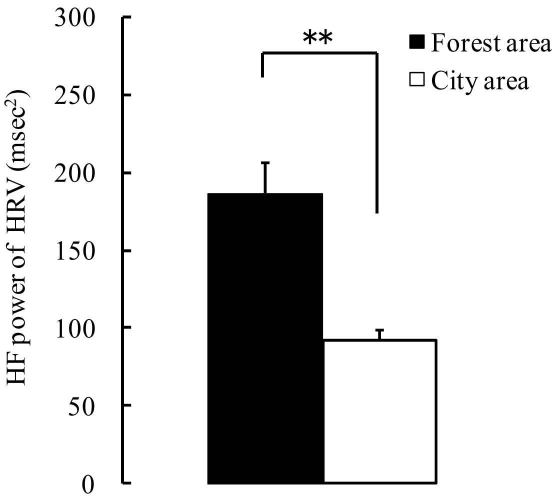
N = 75; Mean ± SD; * $p < 0.05$; p-value by t-test. (Source: Park et al., 2010)

Fig. 6. Change in systolic blood pressure after forest walking.



N = 387; Mean ± SE; ** $p < 0.01$; p-value by t-test. (Source: Park et al., 2011)

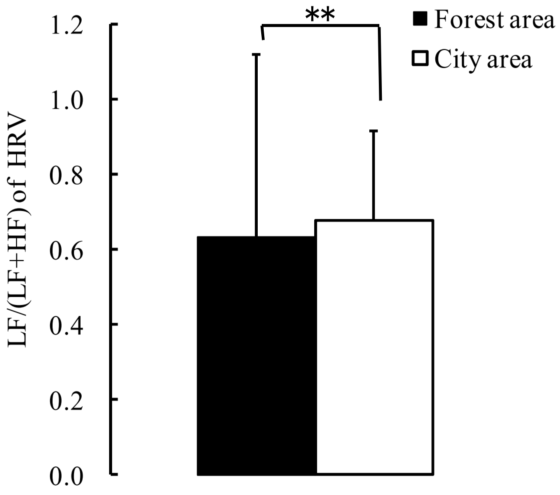
Fig. 7. Change in HF power of HRV on forest viewing.



N = 322; Mean ± SE; ** $p < 0.01$; p-value by t-test. (Source: Park et al., 2011)

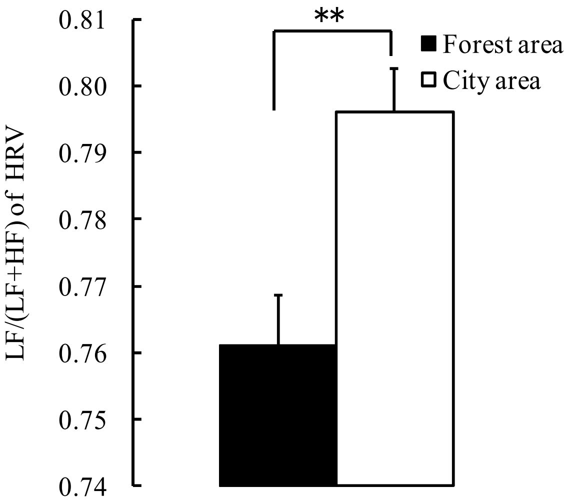
Fig. 8. Change in HF power of HRV on forest walking.

In the results of the HRV analysis, the average power of the HF components increased significantly in the forest environment (viewing and walking) compared with the city environment (Figs. 7 and 8). HF power can be a general indication of parasympathetic nervous activity, which increases when the body is relaxed.



N = 387; Mean ± SE; ** $p < 0.01$; p-value by t-test. (Source: Park et al., 2011)

Fig. 9. Change in LF/(LF + HF) of HRV on forest viewing.



N = 322; Mean ± SE; ** p < 0.01; p-value by t-test. (Source: Park et al., 2011)

Fig. 10. Change in LF/(LF + HF) of HRV on forest walking.

In addition, the average LF/(LF + HF) ratio of HRV, which increases during stress, showed significant differences between the forest and city environments. The average LF/(LF + HF) ratio decreased 7.0% when subjects were viewing the forest landscape and 4.4% when they were walking forest paths, compared with the ratio in the city environment (Figs. 9 and 10). This ratio is an index of sympathetic nervous activity associated with lower stress levels. These results indicate that exposure to the forest environment can reduce stress and induce relaxation.

All indices were generally in excellent agreement. They imply that the forest environment has relaxing and stress-relieving effects on humans. The results also accord with the belief of many people–that forest environments enhance physical relaxation.

Our physiological data are important: because they help explain the mechanism of the relaxation effects of the natural environment. The stress response is mediated by two broad components: the sympathetic–adrenal–medullary (SAM) axis and the HPA axis (Dinan, 2004). The SAM axis is involved in immediate sympathetic activation, which prepares an individual to deal with a stressor and results in changes such as increased heart rate and blood pressure (Vente et al., 2003). Cortisol is released by the HPA axis in response to stress (Seplaki et al., 2004). In our study, while people watched forest landscapes or walked around in the forest, their pulse rate, blood pressure, and cortisol concentration decreased. This supports the idea that the forest environment affects both the main components of the stress response system.

4.1.2 Nature and human immune function

Cortisol concentration also has great significance for human immunological activity (DeAmici et al., 2000). Our previous studies demonstrate the response characteristics of immune function and how immunity can be improved by contact with forest environments. It is well known that the immune system including NK cells plays an important role in defense against bacteria, viruses, and tumors and that stress inhibits immune function (Li et al., 2005ab). Forest environments may reduce stress (Park et al., 2010). Therefore, we speculate that forest environments may have beneficial effects on immune function by reducing stress.

4.1.2.1 Effect of forest environments on NK activity in male subjects

As mentioned above, since 2005, a series of investigations to study the effect of forest environments on human immune function have been conducted in Japan (Li et al., 2007a, 2008a, 2008b, 2010). In the first study (Li et al., 2007a), 12 healthy male subjects aged 37–55 years participated in a 3-day/2-night trip to forest areas at Iiyama, Nagano Prefecture, in northwest Japan in early September 2005. Blood was sampled on the second and third days. NK activity, NK cells, and the expression of GRN, perforin, and GrA/B in lymphocyte were measured. As a control, the same measurements were performed before the trip on a normal working day as a control.

Walking in forests significantly increased NK activity and the number of NK cells (Li et al., 2007a). It has been reported that NK cells kill tumor or virus-infected cells by releasing perforin, granzymes, and GRN via the granule exocytosis pathway (Okada et al., 2003; Li et al., 2004, 2005b). To explore the mechanism of enhancement of NK activity, the effect of being in the forest environment on intracellular levels of perforin, GRN, and GrA/B in PBL was investigated. It was found that forest environments also significantly increased the numbers of intracellular perforin, GRN, and GrA/B-expressing lymphocytes. Taken together, these findings indicate that forest environments can increase NK activity and that this effect might be at least partially mediated by an increased number of NK cells and by induction of intracellular perforin, GRN, and GrA/B (Li et al., 2007a).

However, will a trip to an unforested area (e.g., a city) also increase NK activity? To determine whether taking a city trip can also affect NK activity, we performed another experiment. Eleven healthy male subjects, aged 35–56 years, went on a 3-day/2-night trip to Nagoya City in mid-May 2006 (Li et al., 2008a). The city trip did not increase NK activity, numbers of NK cells, or the expression of the abovementioned anticancer proteins in lymphocyte, indicating that increased NK activity during the forest trip was not due to the trip itself but due to the forest environment (Li et al., 2008a). The important finding is that visiting a forest, rather than a city, increases NK activity and the intracellular levels of perforin, GRN, and GrA/B.

4.1.2.2 How long does the increased NK activity last after a forest park visiting?

Another question remain to be resolved: how long does the increased NK activity last after a forest bathing trip? An investigation was conducted to determine the duration of NK activity after a forest visit (Li et al., 2008a). Twelve healthy male subjects aged 35–56 years participated in a 3-day/2-night trip to forest areas at Agematsu Town, Nagano Prefecture, in northwest Japan in early September 2006 (Li et al., 2008a). NK activity significantly increased in these subjects during the forest visit. NK activity, the numbers of NK cells, and the percentages of GRN-, perforin-, and GrA/B-expressing cells in PBL, which confirmed the previous findings (Li et al., 2007a). The increased NK activity, number of NK cells, and percentages of GRN-, perforin-, and GrA/B-expressing cells lasted more than 7 days and even for 30 days in the cases of NK activity, the number of NK cells, and GRN- and GrB-expressing cells. These findings indicate that a forest bathing trip increased NK activity, the number of NK cells, and the levels of intracellular perforin, GRN, and GrA/B, and that these effects lasted for at least 7 days after the trip (Li et al., 2008a).

4.1.2.3 Effect of forest environments on NK activity in female subjects

Although it has been demonstrated that forest bathing trips enhance human NK activity in male subjects, it still remained to be resolved whether or not forest bathing trips also increase NK activity in female subjects. Since it has been reported that menstruation significantly affects NK activity (Souza et al., 2001),the influence of menstruation on NK activity was controlled during our experiments in female subjects.

In this part of the study (Li et al., 2008b), 13 healthy nurses aged 25–43 years with 4–18 years professional careers participated in a 3-day/2-night trip to forest areas around Shinano Town, Nagano Prefecture, in early September 2007. The trip significantly increased NK activity and the positive rates of NK, perforin-, GRN-, and GrA/B-expressing cells. The increased NK activity and the positive rates of NK, perforin, GRN, and GrA/B-expressing cells lasted for more than 7 days after the trip (Li et al., 2008b), similar to the findings in male subjects (Li et al., 2008a). These findings indicate that a forest bathing trip also increased NK activity, the number of NK cells, and the levels of intracellular anti-cancer proteins in female subjects, and that this effect lasted for at least 7 days after the trip.

4.1.2.4 A day trip to a forest park also increased human NK activity

Although longer trips to forest areas significantly increased NK activity, it was unclear whether a shorter trip to a suburban forest park would also have a similar effect. Our investigation found that a day trip to a forested park also increased NK activity and expression of anticancer proteins in male subjects (Li et al., 2010). In this study, 12 healthy male subjects aged 35–53 years participated in a day trip to a forested park in the suburbs of Tokyo. NK activity and numbers of NK cells, perforin, GRN, and GrA/B-expressing lymphocytes significantly increased, while concentrations of cortisol in blood and adrenaline in urine significantly decreased. The increased NK activity lasted for 7 days after the trip. These findings indicate that the day trip to the forest park also increased the NK activity, number of NK cells, and levels of intracellular anti-cancer proteins, and that this effect lasted for at least 7 days after the trip.

4.1.2.5 Mechanism of increase in NK activity

The question arises as to why does NK activity increase and what factors in the forest environment activate NK cells. We speculate that aromatic volatile substances (phytoncides) derived from trees such as α-pinene and limonene play an important role. We detected several phytoncides such as isoprene, α-pinene, β-pinene, and d-limonene in the forest areas during various trips (Li et al., 2007a, 2008a, 2008b, 2010).

To investigate the effect of phytoncides on NK function, human NK cells were incubated in the presence of phytoncides extracted from trees in the areas visited. NK activity and intracellular levels of perforin, GrA, and GRN were then measured. Phytoncides significantly increased NK activity and intracellular levels of perforin, GrA, and GRN in vitro (Li et al., 2006). Moreover, we found that in vivo exposure to phytoncides from *Chamaecyparis obtusa* stem oil for 3 nights significantly increased NK activity and the percentages of NK cells, perforin, GRN, and GrA/B-expressing cells (Li et al., 2009). These findings suggest that phytoncides contribute to the enhanced NK activity during the forest visits (Li et al., 2006, 2009).

We also found that forest visits significantly decreased the concentrations of adrenaline and noradrenaline in urine (Li et al., 2008a, 2008b, 2010, 2011), which also contribute to increase in NK activity.

Finally, we found that Japanese people living in areas with lower forest coverage had significantly higher standardized mortality ratios for cancers compared with people living in areas with higher forest coverage, suggesting that forest environments may partially contribute to decreased mortality ratios for some cancers (Li et al., 2008c).

Taken together, these findings indicate that forest visits increase NK activity, which was mediated by increases in the number of NK cells and the levels of intracellular anti-cancer proteins. Phytoncides released from trees as well as decreased production of stress hormones may also partially contribute to this increased NK activity (Li, 2010a; Li & Kawada, 2011). Because NK cells can kill tumor cells by releasing anti-cancer proteins such as perforin, GRN, and GrA/B (Okada et al., 2003; Li et al., 2004, 2005b) and forest visits increase NK activity and intracellular levels of anticancer proteins, we can conclude that forest visits may have a preventive effect on cancer cell generation and development.

4.2 Indoor experiments

To clarify the relaxation effects of nature therapy, studies must be conducted both in the field and indoors for comparison purposes. A field study is very valuable in revealing the effects of nature therapy, but reproducibility cannot be assured because of the ever-changing conditions in a field environment. On the other hand, reproducibility can be achieved in indoor studies, where different stimuli can be used and the relaxation effects induced by each stimulus can be examined. Another advantage of indoor studies is that more detailed measurements can be taken, allowing us to focus on the physiological mechanisms leading to a relaxed state (Sakuragawa et al., 2005, 2008; Tsunetsugu & Miyazaki, 2005; Tsunetsugu et al., 2002, 2005, 2007a, 2010).

4.2.1 Olfactory stimulation study

4.2.1.1 Phytoncides

The term “phytoncide” is derived from “phyto,” meaning plant, and “cide,” meaning killing. This term was first used by B. P. Tokin in the Soviet Union around 1930. Later, in 1942, Tokin wrote an article in a booklet published by the National Medical Publishing House in Moscow. In 1946, the term appeared in the first issue of the *Journal of Clinical and Experimental Medicine* (*Igaku No Ayumi*) published in Japan. As stated in these articles, Tokin originally considered that phytoncides were volatile ingredients of plant oils. However, in his later publication, “A Mysterious Phytoncide in Plants” issued in 1980, he defined phytoncides as the “substances produced by all kinds of plants, which may or may not be volatile and which have an influence on other organisms.” The major substances produced by forests and lumber, such as α-pinene and limonene, are good examples of phytoncides. The strong smell of onion or garlic can stimulate our lacrimal (tear) glands when cooking because these smells also contain phytoncides.

4.2.1.2 Relaxation effects of phytoncides

When the air in a forest is analyzed, more than 100 different types of phytoncides can be detected. In many cases, α-pinene and limonene are the major components. To clarify the physiological effects of relaxation uniquely induced by phytoncides, we conducted an inhalation study using α-pinene and limonene in an indoor artificial climate room. Test subjects were exposed to a low concentration of each substance, and blood pressure readings were taken once per second for a period of 90 s. Statistically significant reductions in systolic blood pressure were noted after inhalation of α-pinene and limonene, respectively.

We also conducted an inhalation study using fragrances produced by the wood chips of *sugi* (*Cryptomeria japonica*) and *hiba* or *asunaro* (*Thujopsis dolabrata*). We found that systolic blood pressure decreased significantly after inhalation of these fragrances. Brain activity was also significantly subdued. Subjective assessments using a questionnaire indicated that the subjects were in a state of natural comfort. We can therefore interpret these data as indicating that inhalation of the fragrances of wood chips of *sugi* and *hiba* has a relaxing effect on humans. Interestingly, even in those volunteers who found the fragrances of *sugi* and *hiba* unpleasant, systolic blood pressure did not increase and no stressful conditions were observed.

As mentioned previously, human biological functions are naturally synchronized with the rhythms of the environment. This explains why even when our subjects disliked the fragrance of *sugi*, they still did not experience stress because of the inherent natural adaptation of the human biological system. The same phenomena have been observed in our studies involving tactile and visual stimulation.

4.2.2 Visual stimulation study

Using the indoor artificial climate room, we have measured the physiological responses of volunteers looking at all kinds of views as they walked in the forest or took part in forest therapy. We found that changes elicited by nature-derived visual stimulation were very similar among all subjects, as reflected by subdued activity in the cerebral prefrontal cortex and the autonomic nervous system and reduced blood pressure. Even in cases of single sensory stimulation, the human body demonstrated the ability to find its natural state: a relaxed condition. The only exception to this was when the subjects were viewing *sakura* (Japanese cherry trees). When the volunteers saw cherry trees in full bloom, elevated activity of the cerebral prefrontal cortex and increased heart rate were observed, implying body excitement. We regard this as a case where the fascinating natural beauty of *sakura* could actually change the physiological state of the subjects of our study.

4.2.3 Auditory stimulation study

Auditory stimulation in forests can be a factor in physiological relaxation. Therefore, activity in the prefrontal area of the brain and HRV during auditory stimulation were investigated. Subjects were asked to listen to various sounds from the forest, ranging from the noise of a stream to the singing of nightingales and other birds with their eyes closed. As with the visual stimulation study, we observed subdued activity in the prefrontal cortex and the sympathetic nervous system, indicators of the physiological effects of exposure to these sounds. However, subjects who imagined being in the forest during the experiment showed more signs of relaxation, while those who did not have much interest in the sound or who associated the sound of the forest stream with the flushing sound of a toilet exhibited no objective relaxation effects. Thus, the same sound may be interpreted in different ways by different people, and its effects may therefore differ.

5. Individual differences in health benefits

5.1 Individual differences in human responses

People respond differently even to the same stimuli. There is always considerable variation in the response magnitude between individuals and in the direction of the response. In a previous experiment, the authors found that changes in hemoglobin concentrations in the prefrontal area during exposure to the taste and odor of a piece of chocolate varied from person to person (Fig. 11). In some participants, hemoglobin concentrations increased, which implies increased activity in the prefrontal cortex, while in other participants, it decreased, which reflects decrease in brain activity.

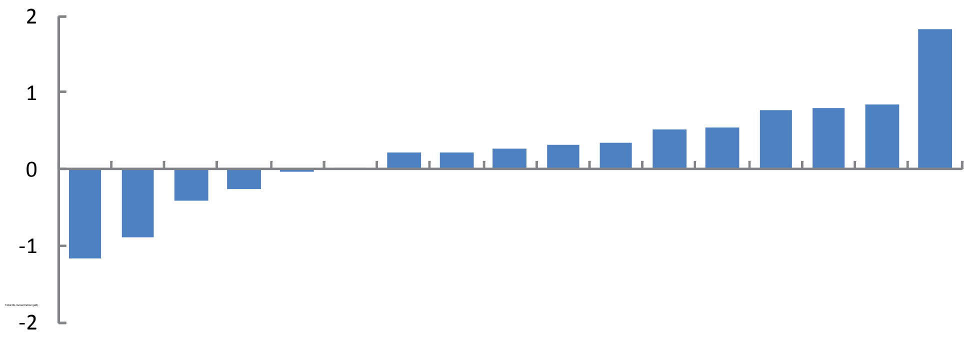


Fig. 11. Changes in total hemoglobin concentration in the left prefrontal rea in 17 young male subjects

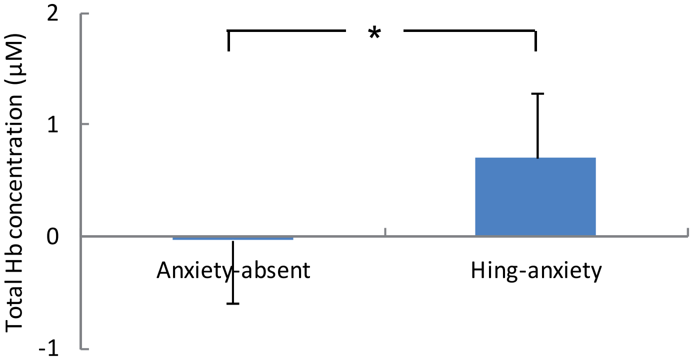
Why do we respond differently? What causes this variation? In an attempt toanswer these questions, Korte et al. (2005) revealed that two types of stress-coping personalities are widely observed in the animal kingdom: aggressive (proactive Hawks) and relatively cooperative (passive Doves). They have different physiological characteristics and show different reactions to changes in their environments. Both types exist in equilibrium within a population and have different strategies to cope with stress, which are effective under different environmental conditions.

Thus, we investigated individual personality as one of the possible factors affecting physiological response (Tsunetsugu et al., 2003). To analyze the previously mentioned hemoglobin data, we examined anxiety to determine if being easily anxious was associated with variations in brain activity. First, we determined each participant's personality type according to their scores on personality tests and divided them into two groups: high anxiety or no anxiety. We then divided the participants into two other groups according to hemoglobin changes, namely the increased and decreased groups, and investigated how many participants of each personality type fell into the two groups. The results were somewhat different than we expected. Significant biases were observed between the increased and decreased groups in terms of the number of high-anxiety people (Table 1). A significant difference was also observed when groups of different anxiety levels were compared in terms of average values of change in hemoglobin concentration (Fig. 12). The high-anxiety group showed a large increase in brain activity during stimulation, while the anxiety-absent group showed only a slight change. These results suggest that individual personality traits could influence physiological responses.

Individual differences in physiological responses have been not scientifically investigated in our study on nature therapy mainly because of problems of accuracy or relevance of the measurement. People nevertheless show large variations and it is naturally necessary to consider this when treating with nature therapy.

	High-anxiety	Anxiety-absent	Total
Decreased group	0	5	5
Increased group	6	6	12
Total	6	11	17

Table 1. Number of participants in each category and cerebral activity in response to olfactory and gustatory stimuli



N = 11 (anxiety-absent), N = 6 (high-anxiety); Mean ± SD; * $p < 0.05$; p-value by unpaired t-test.

Fig. 12. Changes in total hemoglobin concentration in the left prefrontal area of two different personality groups.

Recent improvements in technology related to physiological measurements have enabled data to be obtained from more participants with relative ease. Consequently, discussion on how to approach individual variations scientifically will be ongoing in the near future. In the following section, some examples of analysis focusing on individual personality and the initial values of physiological parameters will be presented.

5.2 Personality and physiological responses

As mentioned previously, certain personality traits may explain individual variations in cerebral activity during olfactory and gustatory stimulation. Here we discuss whether personality traits can also explain individual variations in physiological responses to natural environments.

	Morning	Walking (N = 43)			Viewing (N = 115–116)		
	(N = 117)	Before	After	Change	Before	After	Change
Type A	0.238**	0.282$^+$	0.372*	0.323*	0.260**	0.276**	0.043
Anxiety	0.125	0.158	0.255$^+$	0.288$^+$	0.156$^+$	0.204*	0.115

**p < 0.05, *p < 0.01. "Change" values were calculated by subtracting the "Before" value from the "After" value.

Table 2. Correlation coefficients between personality scores and systolic blood pressure

Experiments were performed in 10 forests in Japan from May to July in 2005, 2006, and 2007. Subjects were 120 male students (12 male students at each experimental site) aged 21.9 ± 1.6 years. Scores for type A behavior pattern and trait anxiety for each subject were calculated according to KG's Daily Life Questionnaire and the Japanese version of the State-Trait Anxiety Inventory (STAI). Systolic blood pressure was measured five times: in the morning before breakfast at the place of accommodation, before and after the subjects walked a predetermined course in the forest for 14 ± 2 min, and before and after they sat still on a chair viewing the scenery in the forest for 14 ± 2 min. Correlation coefficients were calculated between scores of the personality questionnaires and blood pressure (Table 2).

There was a significant positive correlation between type A scores and systolic blood pressure measured in the morning before breakfast, after walking, before watching, and after watching. A significant positive correlation was also observed between type A scores and changes in systolic blood pressure due to walking, which indicated that systolic blood pressure increased in type A subjects while it decreased in type B subjects (Fig. 13). Trait anxiety and systolic blood pressure showed a weak positive correlation after walking and before watching and a significant positive correlation after watching.

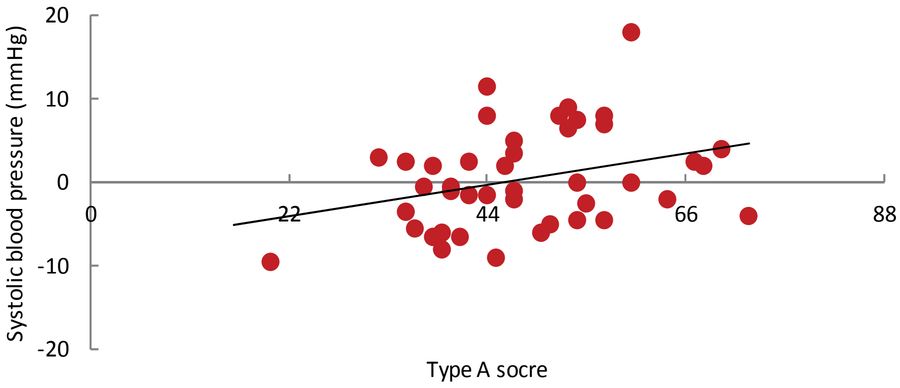


Fig. 13. Relationship between type A scores and changes in systolic blood pressure after walking approximately 14 min in a forest environment (N = 43).

In another study, we focused on salivary alpha-amylase (sAA), a relatively new index of sympathetic nervous activity. Type A behavior patterns and trait anxiety were investigated to determine their effect on sAA levels. sAA of 82 male participants (aged 22.2 ± 1.6 years) was measured four times during their time in a forest environment as well as in the morning and evening. Comparing the mean values of the type A and B groups, we found that the morning sAA levels were significantly lower in the type A group than in the type B group (Fig. 14), whereas morning salivary cortisol levels did not differ between the two groups. The difference in sAA levels between the two groups was maintained for the entire day(Fig. 14).

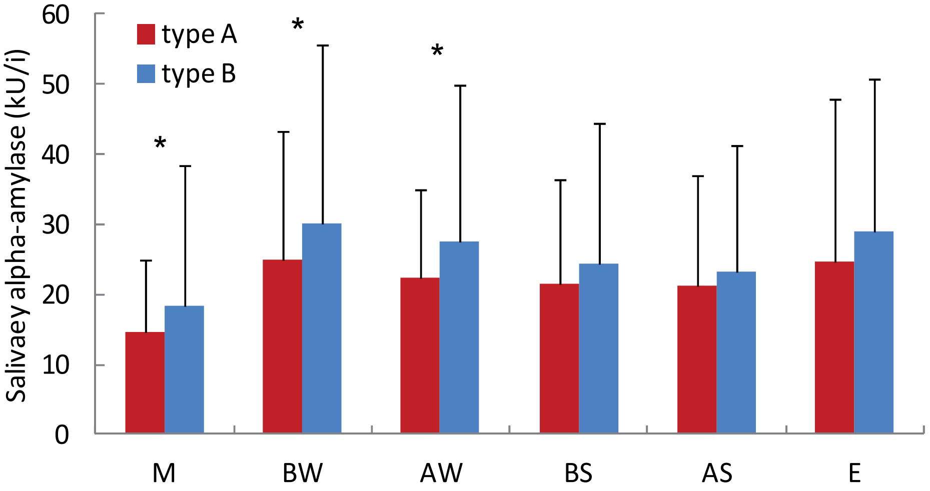


Fig. 14. Time-course changes in salivary alpha-amylase activity in type A and type B groups (mean ± SD, N = 52 (type A), N = 30 (type B). $^*p < 0.05$ (unpaired t-test)).

These results demonstrated that personality traits such as a type A behavior pattern or trait anxiety could be among the factors causing individual variations in the baseline values (values in the morning or before walking and before viewing) of certain physiological parameters and could thus affect changes in physiological responses to forest therapy.

5.3 Baseline values and physiological responses to forest therapy

The law of initial value originally propounded by Wilder (1957) states that the intensity and direction of a body function depend largely on the initial level of that function. Wilder demonstrated that blood pressure or pulse rate responses were higher when the initial values were lower. He also pointed out that most investigators failed to consider the initial values. Does the law of initial value hold true in the case of rather new physiological indicators such as salivary biomarkers?

Stress levels may be assessed by determining the amounts of immune substances secreted in response to stressors. Immunoglobulin A (IgA), an antibody found in blood and other body fluids, is often considered an index of immune system activity. IgA concentration is known to decrease under severe chronic stress (Martin & Dobbin, 1988), whereas it has been reported to increase in response to acute stress (Yamada et al., 1995). Interestingly, however, it has also been demonstrated that IgA levels in those engaged in stressful jobs are significantly higher than in those not engaged in comparable activity (Henningsen et al., 1992). In our field studies, the results for IgA concentration were inconsistent. One of the reasons for this ambiguity could be the large individual differences in baseline values of salivary IgA.

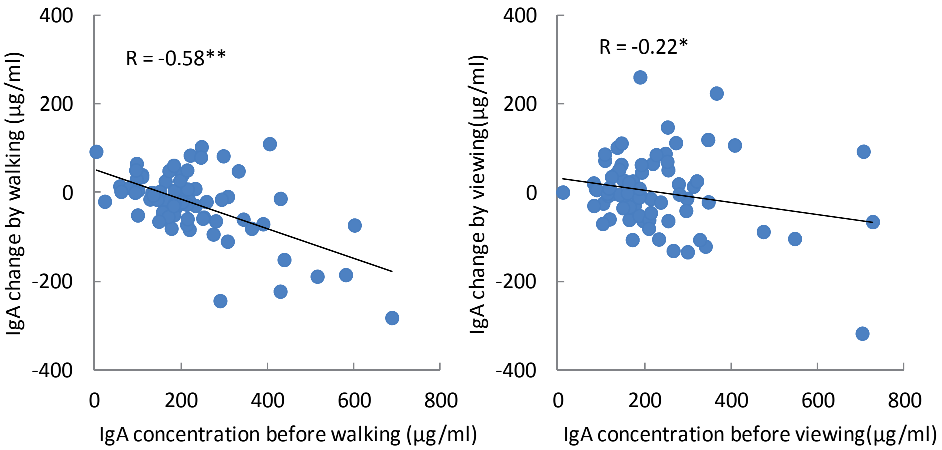


Fig. 15. Relationship between the initial values of salivary IgA concentration and relative changes when walking or viewing in forest environments (N = 75, $^{**}p < 0.01$, $^{*}p < 0.05$).

We investigated salivary IgA concentration in 84 male subjects (22.2 ± 1.6 years old) taken when they walked along a predetermined course for 10–15 min and also when they were seated watching the landscape for 10–15 min in a forest environment. Saliva was sampled four times, i.e., before and after the walk and before and after the landscape viewing. Correlation analysis was performed between the initial values (before walking/viewing) and the changes (after walking/viewing minus before walking/viewing).

A significant negative correlation between the initial value and the degree of change when walking and watching was found (Fig. 15). These results support Wilder's law of initial value and suggest that it is necessary to take the initial value into account in order to understand salivary biomarker data.

6. Current initiatives on nature and human health

The relationship between human health and well-being and forest ecosystems has in recent years received increasing attention not only within multidisciplinary research but also in international and national discussions and policy processes. One global initiative is the International Union of Forest Research Organizations (IUFRO) special project (task force) on Forests and Human Health (IUFRO, 2011). The particular aim of this task force is to specifically address the relationship between forests and human health on a global scale. Its purpose is to support the cross-sectoral dialogue between the different players in this field, especially forestry and health professionals.

The IUFRO Task Force is coordinated by the Finnish Forest Research Institute. It has two priorities: maximizing health benefits of forests and managing health risks connected with forests. The activities of the Health Task Force include organizing conferences and round-table discussions, producing a state-of-the-art report, releasing newsletters and other publications, delivering information, and improving networking through web portals. The task force also aims to facilitate new international research projects and to promote the health benefits of forests. It has conducted several international symposiums and events since 2007.

Another global effort is the Cooperation on Health and Biodiversity (COHAB) initiative, which responds to gaps in awareness and policies linking biodiversity with human health and well-being (COHAB, 2011). This initiative aims to establish a framework to support existing activities on international development, biodiversity conservation, and population health and to support the United Nation's Millennium Development Goals.

Another recent European example of a joint research effort is the scientific and technical network of Forests, Trees, and Human Health and Well-Being (COST action E39) funded by the European Commission. This network summarized current knowledge about the contribution of natural places to the health and well-being of people in Europe and identified future research needs (Nilsson et al., 2010).

Practical organizations such as the International Union for Conservation of Nature (IUCN) have also launched global initiatives to raise awareness of the value of nature areas to human health. IUCN has established the Healthy Parks Healthy People (HPHP) Task Force, which is coordinated by Parks Victoria in Australia. This action follows on from the highly successful first International Healthy Parks Healthy People Congress held in Melbourne in April 2010. The HPHP Task Force aims to provide guidance to IUCN, national policymakers, and practitioners with regard to the relationships between human health, community well-being, economic prosperity (for example, nature-based tourism), ecosystem services, and protected areas. It also seeks to establish alliances with government and nongovernment sectors to progress research and development of the links between parks, society, and economies and to demonstrate the health benefits of nature (Healthy Parks, 2011). The link between health benefits and nature conservation areas such as national parks has recently been actively incorporated into policy documents and management goals of protected areas in Australia, Canada, and Finland, for example.

It seems evident that forest-based health-related products and services have considerable future potential. There is also an increasing understanding that wider application of forest therapy and provision of easily accessible nature-based health services could significantly reduce public health care budgets. However, putting research into practice calls for stronger cooperation between different sectors, especially between forest, health, and environmental professionals (Karjalainen et al., 2010).

A stronger dialogue between researchers in different disciplines will improve theoretical frameworks and methodology. Moreover, a firmer evidence base is needed on the physiological and psychological health benefits of forest therapy. For example, the influence of the quality of forest environments on the therapeutic and restorative benefits of forests is unknown. Cultural, individual, and social differences in experiencing the health benefits of green environments are also not fully understood (Nilsson et al., 2010). In forest-based therapy, a more solid scientific basis for practical applications is also required. Future research might investigate the long-term health benefits of exposure to nearby nature in residential or working environments. These and other questions are currently addressed in recently launched studies such as "Stress reducing effects of urban green areas"' funded by the Academy of Finland and Japan Society for the Promotion of Science (2011–2012).

Although there are promising examples of implementing research results in practice, major implementation of these results is still lacking. An important issue is how to assess the value of the benefits resulting from nature care and other "green-health" approaches. This requires the development of valuation systems to quantify the costs and benefits of interventions (Willis & Crabtree, 2010). At present, public funds are still the main source of direct and indirect payments for health care services. Without an overview of the economic dimensions of these nature-based health-promoting activities, it will be difficult to effectively promote them in a policy setting.

Today, lifestyle-related health problems are of concern in all developed countries. Health problems increasingly relate to modern lifestyles, which are more sedentary and stressful, and mainly oriented indoors. Forests in urban and rural areas have great potential for promoting healthier lifestyles and improved mental health in urbanized societies. However, public authorities do not yet widely promote the use of forests and nature for improved health. Natural areas are still not considered a necessity, particularly in urban areas where the competition for land is intense and land values are high. Compact city policies have led to even greater pressure on urban green areas (Tyrväinen et al., 2005). The health benefits of nature could be better incorporated into urban and land-use planning or in discussions about the need for urban densification. Therefore, the public health benefits of forests must be better understood and more effectively communicated.

7. Conclusion

In this chapter, we introduced nature therapy from the perspective of preventive medicine, illustrating the relationship between nature and human beings, describing the concepts of comfort and health, providing concrete methodology for investigating the effects of nature therapy, and outlining the results of our scientific investigations of the effects of forest therapy on physiological relaxation and immune function. In addition, we also addressed the issues of individual differences in responses to the therapeutic effects of nature. However, despite our efforts, more scientific evidence is required to verify the effects of nature therapy on humans. In future, data from a range of subjects, including young children and middle-aged and elderly people, will facilitate study of the preventive medical effects of nature.

8. Acknowledgments

We are grateful to Prof. and Dr. Tomoyuki Kawada and the other staff at the Department of Hygiene and Public Health, Nippon Medical School.

9. References

Brod, C. (1984).*Technostress: The human cost of the computer revolution*. MA: Addison-Wesley, ISBN 0-201-11211-6.

Cacioppo, J.T., Berntson, G.G., Binkley, P.F., Quigley, K.S., Uchino, B.N., & Fieldstone, A. (1994). Autonomic cardiac control. II. Noninvasive indices and basal response as revealed by autonomic blockades. *Psychophysiol*, Vol.31, pp. 586-598, ISSN 0048-5772

COHAB Initiative. Co-Operation On Health And Biodiversity. October 6, 2011, Available from: http://www.cohabnet.org/

COST Action E39. Forests, Trees and Human Health and Wellbeing. October 6. 2011, Available from: <http://www.e39.ee/>

De Amici, D., Gasparoni, A., Chirico, G., Ceriana, P., Bartoli, A., Ramajoli, I. & Moretta, A. (2000). Natural killer cell activity and delivery: Possible influence of cortisol and anesthetic agents. A study on newborn cord blood. *Biol Neonate*,Vol.78, pp. 70–72, ISSN 0006-3126

Dinan, T.G. (2004). Stress and the genesis of diabetes mellitus in schizophrenia. Brit J psychiat, Vol. 184, pp. 72–75, ISSN 0007-1250

Healthy Parks Healthy People Task Force. October 6. 2011, Available from: <http://www.hphpcentral.com/>

Henningsen, G.M., Hurrell, J.J. Jr., Baker, F., Douglas, C., MacKenzie, B.A., Robertson, S.K., & Phipps, F.C. (1992). Measurement of salivary immunoglobulin A as an immunologic biomarker of job stress. *Scand J Work Environ Health*, Vol.18, Suppl.No. 2, pp. 133-136, ISSN0355-3140

IUFRO Task Force on Forests and Human Health. October 6. 2011. Available from: http://www.forhealth.fi

Karjalainen, E., Sarjala, T. & Raitio, H. (2010). Promoting human health through forests: overview and major challenges. *Environ Health Prev Med*, Vol., No.1, pp. 1-8, ISSN 1342-078X

Kirschbaum, C. & Hellhammer, D.H. (1989). Salivary cortisol in psychobiological research: An overview. *Neuropsychobiology*, Vol.22, pp. 150-169, ISSN 0302-282X

Kirschbaum, C. & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology*, Vol.19, pp. 313-333, ISSN 0306-4530

Kobayasi, H., Ishibashi, K., Noguchi, H. (1999). Heart rate variability, an index for monitoring and analyzing human autonomic activities. *J Physiol Anthropol Appl Human Sci*, Vol.18, No.2, pp. 53-59, ISSN 1345-3475

Korte, S.M., Koolhaas, J.M., Wingfield, J.C. & McEwen, B.S. (2005).The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. *Neurosci Biobehav Rev*, Vol.29, No.1, pp. 3-38, ISSN0149-7634

Lee, J., Park, B.J., Tsunetsugu, Y., Kagawa, T. &Miyazaki, Y. (2009). The restorative effects of viewing real forest landscapes: Based on a comparison with urban landscapes. *Scand J Forest Res*, Vol.24, pp. 227-234, ISSN0282-7581

Lee, J., Park, B.J., Tsunetsugu, Y. & Miyazaki, Y. (2011a). Forests and human health - recent trends in Japan. In : *Forest Medicine*, Q. Li, (Ed.), pp. 243-257, Nova Science Publishers, NY. ISBN 978-1-62100-000-6

Lee, J., Park, K.T., Lee, M.S., Park, B.J., Ku, J.H., Lee, J.W. Oh, K.O., An, K.W. & Miyazaki, Y. (2011b). Evidence-based Field Research on Health Benefits of Urban Green Area. *J Korean Insti Landsc Archit*, Vol. 39, No.5, pp. 111-118, ISSN 1225-1755

Lee, J., Park, B.J., Tsunetsugu, Y., Ohira, T., Kagawa, T. & Miyazaki, Y. (2011c). Effect of forest bathing on physiological and psychological responses in young Japanese male subjects. *Public Health*, Vol.125, pp. 93-100, ISSN 0033-3506

Li, Q., Nakadai, A., Takeda, K. & Kawada, T. (2004). Dimethyl 2,2-dichlorovinyl phosphate (DDVP) markedly inhibits activities of natural killer cells, cytotoxic T lymphocytes and lymphokine-activated killer cells via the Fas-ligand/Fas pathway in perforin-knockout (PKO) mice.*Toxicology*, Vol.204, No.1, (November 2004), pp. 41-50, ISSN0300-483X

Li, Q., Liang, Z., Nakadai, A. & Kawada, T. (2005a). Effect of electric foot shock and psychological stress on activities of murine splenic natural killer and lymphokine-activated killer cells, cytotoxic T lymphocytes, natural killer receptors and mRNA transcripts for granzymes and perforin.*Stress*, Vol.8, No.2, (June 2005), pp. 107-116, ISSN 1025-3890

Li, Q., Nakadai, A., Ishizaki, M., Morimoto, K., Ueda, A., Krensky, A.M. & Kawada, T. (2005b). Dimethyl 2,2-dichlorovinyl phosphate (DDVP) markedly decreases the expression of perforin, granzyme A and granulysin in human NK-92CI cell line.*Toxicology*, Vol.213, No. 1-2, (September 2005), pp. 107-116, ISSN0300-483X

Li, Q., Nakadai, A., Matsushima, H., Miyazaki, Y., Krensky, A.M., Kawada, T. & Morimoto, K. (2006). Phytoncides (wood essential oils) induce human natural killer cell activity.*Immunopharmacol Immunotoxicol*, Vol.28, No.2 (April–June 2006), pp. 319-333, ISSN 0892-3973

Li, Q., Kobayashi, M., Wakayama, Y., Inagaki, H., Katsumata, M., Hirata, Y., Hirata, K., Shimizu, T., Kawada, T., Ohira, T., Park, B.J., Kagawa, T. & Miyazaki, Y. (2009). Effect of phytoncide from trees on human natural killer cell function. *Int J Immunopathol Pharmacol*, Vol.22, No. 4, (October-December 2009), pp. 951-959, ISSN 0394-6320.

Li, Q., Morimoto, K., Nakadai, A., Qu, T., Matsushima, H., Katsumata, M., Shimizu, T., Inagaki, H., Hirata, Y., Hirata, K., Kawada, T., Lu, Y., Nakayama, K. & Krensky A.M. (2007b). Healthy lifestyles are associated with higher levels of perforin, granulysin and granzymes A/B-expressing cells in peripheral blood lymphocytes.*Prev Med*, Vol.44, No.2, (February 2007), pp. 117-123, ISSN 0091-7435

Li, Q., Morimoto, K., Kobayashi, M., Inagaki, H., Katsumata, M., Hirata, Y., Hirata, K., Suzuki, H., Li, Y.J., Wakayama, Y., Kawada, T., Park, B.J., Ohira, T., Matsui, N., Kagawa, T., Miyazaki, Y. & Krensky, A.M. (2008a). Visiting a forest, but not a city, increases human natural killer activity and expression of anti-cancer proteins.*Int J Immunopathol Pharmacol*, Vol.21, No. 1, (January–March 2008), pp.117-27, ISSN 0394-6320

Li, Q., Morimoto, K., Kobayashi, M., Inagaki, H., Katsumata, M., Hirata, Y., Hirata, K., Shimizu, T.,Li, Y.J., Wakayama, Y., Kawada, T.,Ohira, T., Takayama, N.,Kagawa, T. & Miyazaki, Y. (2008b).A forest bathing trip increases human natural killer activity and expression of anti-cancer proteins in female subjects.*J Biol Regul Homeost Agents*, Vol.22, No.1, (January–March 2008), pp. 45-55, ISSN 0393-974X

Li, Q., Kobayashi, M. & Kawada, T. (2008c). Relationships between percentage of forest coverage and standardized mortality ratios (SMR) of cancers in all prefectures in Japan. *The Open Public Health Journal*, Vol.1, No.1, pp. 1-7, ISSN 1874-9445

Li, Q., Kobayashi, M., Wakayama, Y., Inagaki, H., Katsumata, M., Hirata, Y., Hirata, K., Shimizu, T., Kawada, T., Ohira, T., Park, B.J., Kagawa, T. &Miyazaki, Y. (2009). Effect of phytoncide from trees on human natural killer cell function. *Int J Immunopathol Pharmacol*, Vol.22, pp. 951-959, ISSN 0394-6320.

Li, Q. & Kawada, T. (2011). Effect of forest environments on human natural killer (NK) activity.*Int J Immunopathol Pharmacol*, Vol.24, No.1 Suppl, (January–March 2008), pp. 39S-44S, ISSN 0394-6320

Li, Q., Kobayashi, M. ,Wakayama, Y., Inagaki, H., Katsumata, M., Hirata, Y., Hirata, K., Shimizu, T., Kawada, T., Ohira, T., Park, B.J., Kagawa, T. & Miyazaki, Y. (Effect of phytoncide from trees on human natural killer function. *Int J Immunopathol Pharmacol*, Vol., ISSN 0394-6320

Li, Q., Kobayashi, M., Inagaki, H., Hirata, Y., Li, Y.J., Hirata, K., Shimizu, T., Suzuki, H., Katsumata, M., Wakayama, Y., Kawada, T., Ohira, T., Matsui, N. & Kagawa T. (2010). A day trip to a forest park increases human natural killer activity and the expression of anti-cancer proteins in male subjects.*J Biol Regul Homeost Agents*, Vol.24, No.2, (April–June 2010), pp. 157-165, ISSN 0393-974X

Li, Q. (2010a). Effect of forest bathing trips on human immune function. *Environ Health Prev Med*, Vol.15, No.1, (January 2010), pp. 9-17, ISSN 1342-078X

Li, Q. (2010b). NK cell assays in immunotoxicity testing.*Methods Mol Biol*Vol.598, pp. 207-219,ISSN 1064-3745

Li, Q., Otsuka, T., Kobayashi, M., Wakayama, Y., Inagaki, H., Katsumata, M., Hirata, Y., Li, Y., Hirata, K., Shimizu, T., Suzuki, H., Kawada, T & Kagawa, T. (2011). Acute effects of walking in forest environments on cardiovascular and metabolic parameters. *Eur J Appl Physiol*, Vol.111, No.11, (November 2011), pp.2845-2853, ISSN 1439-6319

Martin, R.A. & Dobbin, J.P. (1988). Sense of humor, hassles, and immunoglobulin A. evidence for a stress-moderating effect of humor. *Int J Psychiatry Med*, Vol.18, No.2, pp. 93-105, ISSN0091-2174

Matsunaga, K., Park, B.J., Kobayash, H. &Miyazaki, Y. (2011). Physiologically Relaxing Effect of a Hospital Rooftop Forest on Elderly Women Requiring Care. *J Am Geriatr Soc*, in press,ISSN0002-8614

Mishima, R., Kudo, T., Tsunetsugu, Y., Miyazaki, Y., Yamamura, C. & Yamada, Y. (2004). Effects of sounds generated by a dental turbine and a stream on regional cerebral blood flow and cardiovascular responses. *Odontol*, Vol.92, pp. 54-60, ISSN1618-1247

Miyazaki, Y., Park, B.J. & Lee J. (2011). Nature therapy. In: *Designing our future: Local Perspectives on bioproduction, ecosystems and humanity* (Sustainability Science Vol. 4), Osaki, M., Braimoh, A., Nakagami, K., (Ed.), pp. 407-412, United Nations University Press, Tokyo, ISBN 978-92-808-1183-4

Nilsson, K., Sangster, M., Gallis, C., Hartig, T., de Vries, S., Seeland, K. & Schipperin. J. (2010). *Forests, Trees and Human Health*. Springer Science+Business Media, New York, Dordrecht, Heidelberg, London. 409 p. + Appendices. ISBN 978-9048198054

Okada, S., Li, Q.,Whitin, J.C., Clayberger, C. & Krensky, A.M. (2003). Intracellular mediators of granulysin-induced cell death.*J Immunol*, Vol.171, No.5, (September 2003), pp. 2556-2562, ISSN 0022-1767

Park, B.J., Tsunetsugu, Y., Kasetani, T., Hirano, H., Kagawa, T., Sato, M. & Miyazaki Y. (2007). Physiological effects of Shinrin-yoku(taking in the atmosphere of the forest)

-using salivary cortisol and cerebral activity as indicators. *J Physiological Anthropol*, Vol.26, pp. 123-128, ISSN 1880-6791

Park, B.J. & Miyazaki, Y. (2008a). Physiological Effects of Viewing Forest Landscapes - Results of field tests in Atsugi city, Japan. (in Korean). *J Korean Forest Soc*, Vol.97, pp. 634-640, ISSN 0445-4650

Park, B.J., Tsunetsugu, Y., Ishii, H., Furuhashi, S., Hirano, H., Kagawa, T. & Miyazaki Y. (2008b). Physiological effects of Shinrin-yoku (taking in the atmosphere of the forest) in a mixed forest in Shinano Town, Japan. *Scand J Forest Res*, Vol.23, pp. 278-283, ISSN 0282-7581.

Park, B.J., Kasetani, T., Morikawa, T., Tsunetsugu, Y., Kagawa, T. &Miyazaki ,Y. (2009). Physiological effects of forest recreation in a young conifer forest in Hinokage Town, Japan. *Silva Fenn*, Vol.43, pp. 291-301, ISSN 0037-5330

Park, B.J., Tsunetsugu, Y., Kasetani, T., Kagawa, T. & Miyazaki, Y. (2010).The physiological effects of Shinrin-yoku (taking in the forest atmosphere or forest bathing): evidence from field experiments in 24 forests across Japan. *Environ Health Prev Med*, Vol.15, No.1, pp 18-26, ISSN 1342-078X

Park, B.J., Tsunetsugu, Y., Lee, J., Miyazaki, Y., et al. (2011). Effect of the forest environment on physiological relaxation-the results of field tests at 35 sites throughout Japan. In : *Forest Medicine*, Q. Li, (Ed.), pp. 55-65, Nova Science Publishers, NY. ISBN 978-1-62100-000-6

Pruessner, J.C., Hellhammer, D.H. & Kirschbaum, C. (1999). Burnout, perceived stress, and cortisol responses to awakening. *Psychosom Med*, Vol.61, pp. 197-204, ISSN 0033-3174

Sakuragawa, S., Miyazaki, Y., Kaneko, T. & Makita, T. (2005). Influence of wood wall panels on physiological and psychological responses. *J Wood Sci*, Vol.51, pp. 136-140, ISSN 0021-4795

Sakuragawa, S., Kaneko, T. &Miyazaki, Y. (2008). Effects of contact with wood on blood pressure and subjective evaluation. *J.Wood Sci*, Vol.54, No.2, pp. 107-113, ISSN1435-0211

Seplaki, C.L., Goldman, N., Weinstein, M. & Lin, Y.H. (2004). How are biomarkers related to physical and mental well-being? *J Gerontol A Biol Sci Med Sci*, Vol.59, pp. 201-201, ISSN 1079-5006

Sluiter J.K., Frings-Dresen, M.H.W., Meijman, T.F. & van der Beek, A.J. (2000). Reactivity and recovery from different types of work measured by catecholamines and cortisol: a sysmetic literature overview. *Occup Environ Med*, Vol.57, pp. 298-315, ISSN 1351-0711

Souza, S.S., Castro, F.A., Mendonça, H.C., Palma, P.V., Morais, F.R., Ferriani, R.A. & Voltarelli, J.C. (2001). Influence of menstrual cycle on NK activity. *J Reprod Immunol*, Vol.50, No.2, (May 2001), pp. 151-159, ISSN 0165-0378

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation*, Vol.93, No.5, pp. 1043-1065, ISSN 0009-7322

Tsunetsugu, Y. & Miyazaki, Y. (2005a). Measurement of absolute hemoglobin concentrations of prefrontal region by near-infrared time resolved spectroscopy: examples of experiments and prospects. *J Physiol Anthropol Appl Human Sci*, Vol.24, pp. 469-472, ISSN 1345-3475

Tsunetsugu, Y., Miyazaki, Y., Morikawa, T. & Kamwaki, T. (2003). Personality and physiological response: relationship between Type A behavior pattern trait anxiety, gender role perception and variation in cerebral blood flow change- taking chocolate stimulus as an example-. *Japanese Journal of Physiological Anthropology*, Vol.8, No.4, pp. 259-261, (in Japanese), ISSN 1342-3215

Tsunetsugu, Y. &Miyazaki, Y. (2005). Measurement of absolute hemoglobin concentrations of prefrontal region by near-infrared time resolved spectroscopy: examples of experiments and prospects. *J Physiol Anthropol Appl Human Sci,* Vol.24, pp. 469-472,ISSN1345-3475

Tsunetsugu, Y., Miyazaki, Y. & Sato, H. (2005b). Visual effects of interior design in actual-size living rooms on physiological responses. *Building Environ*, Vol.40, pp. 1341-1346, ISSN 0360-1323

Tsunetsugu, Y., Miyazaki, Y.&Sato, H. (2007). The physiological effects in human s induced by the visual stimulation of room interiors with different wood quantities. *J Wood Sci*, Vol.53, pp. 11-16, ISSN1435-0211

Tsunetsugu, Y., Park, B.J., Ishii, H., Hirano, H., Kagawa, T. & Miyazaki, Y. (2007). Physiological effects of "Shinrin-yoku" (taking in the atmosphere of the forest) in an old-growth broadleaf forest in Yamagata prefecture, Japan. *J Physiological Anthropol,*Vol.26, pp. 135-142, ISSN 1880-6791

Tsunetsugu, Y., Park, B.J. &Miyazaki, Y. (2010). Trends in research related to "Shinrin-yoku" (taking in the forest atmosphere or forest bathing) in Japan.*Environ Health Prev Med*, Vol.15, pp. 27-37,ISSN 1342-078X

Tsunetsugu, Y., Park, B.J. & Miyazaki, Y. (2011). Physiological effects of visual, olfactory, auditory, and tactile factors in the forest environment. In : *Forest Medicine*, Q. Li, (Ed.), pp. 169-181, Nova Science Publishers, NY. ISBN 978-1-62100-000-6

Tyrväinen, L., Pauleit, S., Seeland, K. & de Vries, S. (2005). Benefits and uses of urban forests and trees. In: *Urban Forests and Trees in Europe A Reference Book*. Nilsson, K., Randrup, T.B. & Konijnendijk, C.C. (Eds.), Springer Verlag. pp. 81-114. ISBN 978-3540251262

Vente, W.D, Olff, M., Amsterdam, J.G.C.V., Kamphuis, J.H. & Emmelkamp, P.M.G. (2003). Physiological differences between burnout patients and healthy controls: Blood pressure, heart rate, and cortisol responses. *Occup Environ Med*, Vol. 60, pp. i54–i61, ISSN 1351-0711

Weise, F. & Heydenreich, F. (1989). Effects of modified respiratory rhythm on heart rate variability during active orthostatic load. *Biomed Biochim Acta*, Vol.48, pp. 549-556, ISSN 0232-766X

Wilder, J. (1957). The law of initial value in neurology and psychiatry; facts and problems. *J Nerv Ment Dis*, Vol.125, No.1, pp. 73-86, ISSN0022-3018

Willis, K. & Crabtree, B. (2010). Measuring health benefits of green space in economic terms. In: *Forests, Trees and Human Health*, Nilsson, K., Sangster, M., Gallis, C., Hartig, T., de Vries, S., Seeland, K. & Schipperin. J. (Eds.), Springer Science+Business Media, New York, Dordrecht, Heidelberg, London. pp. 375-402. ISBN 978-9048198054

Yamada, F., Miyata, Y., Takenaka, K. & Tanaka, K. (1995). Secretory IgA as an index of stress reactivity. *Bull Osaka Pref Coll Nurs*, Vol.1, No.1, pp. 47-50, (in Japanese), ISSN 13410989

17

How Can the Empowerment Role of Public Health Nurses (PHNs) Be Fostered? A Review of an Exploratory Research Study Conducted in Ireland and Current Evidence

Teresa Cawley
Empowering People and Organisations, Teresa Cawley
Republic of Ireland

1. Introduction

Public Health nursing is complex because it incorporates an expectation of empowerment and advocacy on behalf of the public health nurse (PHN) in three functions: curative engagement, health promotion, and political advocacy (Cawley and Mannix McNamara 2011). Indeed, the mission of public health nursing calls for the health promotion approach (DOH 1997) and PHNs have a unique opportunity in their roles to work in a health promoting way. International trends are towards more population-focused objectives that concentrate on disease prevention and health promotion (Edgecombe 2005).

There is an established strong case for health professionals including PHNs to work to promote health (Who 1986) and empowerment of health professionals is central to health promotion as is evidenced by the priority it receives in the Ottowa Charter (WHO 1986). Added to this is, the evidence that the empowerment for the healthcare professional influences their ability to engage in empowering practices (DOHC 2003), p 13, Fulton, 1997, Kanter 1983 and Ryan et al, 2006). From management's perspectives, enablement of the achievement of the full potential of the health professional is key to the management of an effective health service (DOHC 2003). Therefore, if PHNs are to fulfil their diverse role it is essential they are empowered in order to effectively advocate for themselves and their clients.

Falk-Rafael (2004) and Fulton (1997) also found that nurses are being asked repeatedly to facilitate empowerment of their clients. The WHO (2000) through the Munich Declaration formally reiterated the pivotal role of nurses and midwives as a force for health in society's efforts to tackle the public health challenges of our time. Falk-Rafael, 2005, Stevens and Hall, 1992, Byrd, 1995, Williamson and Drummond, 2000 and McDonald, 2002) all argue that nurses have a moral/ ethical and professional obligation to be involved in socio-political activities that address structural conditions contributing to health inequalities because they see the impact of these conditions every day. Economic constraints and the need for healthcare systems to cut costs have led to a series of measures that influence nursing, including earlier discharge of patients, downsizing of the professional workforce, changes in staff mixes, restructuring of services and decreased support services for patient care (Swihart, 2006)

Labonte noted that many healthcare providers are relatively powerless in their organisations and need to claim legitimacy or power for themselves in order to be effective in their work with less powerful groups external to their organisation (Labonte 1994). When professionals are not granted professional status, they have great difficulty in establishing an empowering contract with their clients because they lack voice in the situation to be able to do so (Schon 1983). A dilemma facing health practitioners with regard to empowerment is that they are expected to engage in processes that are empowering for clients often without consideration as to whether they experience themselves as empowered practitioners (Ryan et al 2006). Nurses need to be empowered to make decisions about their practice (Barden et al 2011)

This chapter discusses the findings from an exploratory study titled "the empowerment approaches used by a sample of public health nurses in one of their child health programmes and the experience of these approaches by a sample of parents" in the context of current evidence. "*The study*"explored empowerment in relation to; Public Health Nurses (PHN) understanding of the concept, PHNs experience of empowerment in their work, the PHN role in relation to client empowerment, how PHNs perceive they empower their clients and how the empowerment of PHNs could be fostered. "*The study*" also explored the experience by clients of feeling empowered/ disempowered and the empowerment approach of PHNs in the context of the child health screening and surveillance programme in the Health Services Executive West Ireland. This study built on research by DoHC (2003) into the empowerment of nurses in Ireland and on research specific to empowerment and the empowerment role of the PHNs, performed by Jackson et al (1996) and Falk-Rafael (2001). "*The study*" findings indicated that PHNs did not perceive that they are empowered in their work and are unaware of this. This was related mainly to, their workloads and a lack of open access to structures of power and opportunity at work. The findings suggest that the empowerment of PHNs influences their engagement in empowering practices. The research suggests that PHNs need open access to organisational structures of empowerment. The study will be referred to in this chapter as *"the study"*.

2. Background

The PHN role in child health is similar to that of health visitor in the United Kingdom and the PHN in the United States and Canada (Cohen & Reutter, 2007). The PHN in Slovenia, Finland, Iceland and Latvia (Scottish Executive) are generalist nurses with responsibility for providing primary, secondary and tertiary care to a variety of groups (Hanafin et al, 2002).

A number of critical factors were used as the framework for *"the study"*. The autonomy of the healthcare professional/ was identified as the foundation to their achievement of potential and empowerment was the fundamental methodological approach. The empowerment role of the PHN is to foster the autonomy of her clients, so that clients can self-care and a political approach to support the empowerment of clients is proposed. The empowerment concept was reviewed by examining power as a concept, which involved exploring the theories of power. Kanter (1977, 1993) structural factors within the work environment have a greater impact on employee work attitudes and behaviour than personal dispositions or socialization experiences. The ability to mobilize others is directly related to one's own level of empowerment (Kanter 1983). Empowerment has been found to be an effective method to advance nurse's job satisfaction, and has been widely discussed

(Laschinger et al 2001). To support the healthcare practitioner's role as health promoter, Lewin and Urmston (2000) suggest that the Health Service needs to facilitate greater empowerment of nurses. Therefore, if PHNs are empowered in their work they are better placed to achieve their full potential and facilitate the empowerment of their clients.

On examination of the literature the author noted that illness has not declined significantly, and that the main causes of mortality have strong lifestyle factors. This is despite increasing education, screening, and vaccination (Naidoo and Wills, 2000). There have been significant changes in the area of primary healthcare, due in part to the ever-increasing needs by the health service users, and due to earlier discharge into the community. The Primary Care Strategy has also had a strong influence. The impact of such growing pressures means an increased belief that healthcare needs must be met in the community. There is also a greater emphasis on self-care, and this has added to the need for healthcare professionals in the community to work in a health promoting way. In fact, nurses are being asked to empower others (English Nursing Board (ENB) 1991).

Empowerment has a strong political dimension and this requires self-reflection, dialogue, action and commitment to change (Carey 2000). It is defined in political, management and organisational development terms. Real empowerment challenges the healthcare practitioner and the system and when asked to question structures of power, the practitioner begins to question their practices (Ryan et al 2006p 51). Empowerment is an ongoing process towards what Maslow (1968) termed as self-actualisation. Freire (1972) contended that the facilitation of empowerment is a moral obligation. The DoHC (1998) recommended that PHNs should be allowed to focus to a greater extent on their health promotion and disease prevention role. However to date, no mechanism or resources have been provided to support this. Empowerment strategies intended to enable individuals, groups and communities to have greater control over decisions affecting their lives are becoming a required part of public health nursing practice (Jackson et al 1996, p 7).

Health visitors and PHNs have an opportunity to meet clients in their homes, in the real world where health choices are made. The author therefore reflected on and a reviewed a sample of the literature related to public health nursing, health promotion and empowerment. The author noted that roles, workloads and the frameworks within which PHNs work in the republic of Ireland shows some variations to that of PHNs worldwide. The literature review highlighted that a number of studies had explored empowerment in the context of public health nursing internationally. Aston et al (2006) conducted research in Canada into how empowerment as an ideology and a practice in teaching and learning was understood and applied by PHNs in health education with child bearing and child rearing families through interviews with mothers and PHNs. Their findings identified a need for further investigation into the required skills of PHNs that enable them to co-create empowering relations.

A nurse's role of caring can be at odds with empowerment through moral and political dilemmas (Malin and Teasdale 1991). When healthcare workers like PHNs have large case loads and their communities have many and varied healthcare needs, these needs must be met. In meeting these needs, the PHN may be left with little time to work to support empowerment.

An approach that combines critical social theory, organisational theory, and social psychological theories of power and Foucault's work on power was identified by "*the study*" as key to the empowerment of the PHN by supporting an environment that creates the conditions that support reflection and raising of PHN self-efficacy. A Health service that ensured structures of opportunity are maintained for PHNs and places increased emphasis on support, feedback and affirmation as part of the structure of power are needed. In addition, leadership of PHNs and leadership skills of PHNs are required. An action theory approach is needed in practice to bridge the knowledge practice gap to foster the empowerment of PHNs.

3. Methodology of original study

"*The study*" exploratory had two phases. Phase 1 comprised of two focus groups with a purposive sample of nine PHNs. The focus groups explored the concept of empowerment, and what empowered the PHNs and the strategies they use to support empowerment of their clients. Data generated by the focus groups was analysed by a model that integrated Burnard's (1991) method of analysing interview transcripts by thematic content analysis with the three-stage method outlined by Kumar (1999) and Skelton (1997) supported by the authors own methodological questions. Phase 2 consisted of the administration of questionnaires to a sample of 107 parents with a yield of 40.2% on the administered questionnaires. The questionnaires examined the experience by parents of the empowerment approach of PHNs at the child health screening and surveillance checks. SPSS (Statistical package for the social sciences) was used to analyse the data generated.

4. Theoretical framework

4.1 Empowerment as a concept

Empowerment is both an interpersonal process through providing resources to enable others to set and reach goals (Hawks and Hromek 1992, Rodwell 1996) and an outcome being translates as autonomous decision making, self-determination and feelings of self worth(Gibson, 1991). Combining both fosters autonomy and participation. Conger and Kanungo (1988) define empowerment as a process of enhancing feelings of self-efficacy amongst organisational members through the identification of conditions that foster powerlessness and through their removal by formal organisational practices and informal techniques of providing efficacy information. "*The study*" this chapter builds on defined empowerment as a process and an outcome in which the ability, self-efficacy and autonomy of the PHN is enabled and strengthened by the Health Service creating an empowering environment, where clients and healthcare professionals have open channels to all the resources they need to support them in achieving their personally defined potentials (Cawley and Mannix McNamara 2011). Therefore, the PHN has what Rodwell 1996 defined as "effectiveness which is capacity, ability and means".

4.2 Critical social theory

Emancipatory research has the intent to challenge inequalities and disrupt the status quo where necessary and Rose and Glass (2008) attest that its role in contemporary nursing practice is significant as it sets out to raise the consciousness of nurses who are located in marginalised and oppressed positions. It has oppression as its central focus and social change as key objective with fostering an ideology based on the belief that knowledge is socially constituted, historically situated and valuationally based (Henderson, 1995, p 58). Paulo Freire popularised this theory in his efforts to liberate. Wittmann-Price (2004) argued that the first step towards creating an emancipatory healthcare environment is by first recognising the existence of oppression and in turn its impact on nursing care delivery. Indeed in critical social theory, nurses may be seen as oppressed (Gilbert 1995, Fulton 1997 and Ryles 1999).This needs to be followed by praxis which is a constant interaction between action for change and reflection on action. Educational Institutes and the Health Service need to facilitate reflection on power as it has influenced their role development from its inception and how the structures and systems within which they work impact on their empowerment. Foucault attests that power is not exclusively legalised power, decision making or technique (1978) and its disposal requires the assumption of some responsibility (Foucault 1994). Illich (1970, p 18) concurs and suggests that freedom and power are determined by a willingness to accept responsibility for the future. PHNs have many sources of power including; knowledge, skills, role, access to information and location in relation to power etc and these bestow responsibility and accountability.

4.3 Organisational theory

Kanter's (1977) theory postulated that access to structures of power and opportunity are related to position and job context. Kanter's (1977) theory proposes that access to structures of power and opportunity have a far greater influence on work performance than personality factors. The structures of opportunity are the opportunities for professional growth. They result in the display by the employee of a set of positive attitudes, such as higher levels of motivation and commitment. Kanter (1977) conceptualises that structures of power require open connections through the channels of the power sources within the system and if all channels are open then workers have the power needed to be productive. The channels of power are:

- Lines of supply or access to resources needed to carry out the job efficiently.
- Lines of information, knowing all one needs to know to do the job.
- Lines of support or feedback on performance and continuing approval of actions.

(Kanter 1977)

Access to empowering structures is facilitated by formal and informal job characteristics. Formal power is enhanced when jobs are flexible, central to the organisational goals and allow employees to exercise creativity and discretionary decision-making. Informal power is derived from the development of effective relationships and communication channels with sponsors, peers, subordinates and cross-functional groups, both within and outside the organisation. The organisational factors that were found to influence PHN empowerment in "*the study*" were; role and workload of the PHN, the PHN management structures, and the Health Service structure. Psychological empowerment is the logical outcome of managerial efforts to create Kanter's structural conditions of empowerment (Laschinger 2001, p 11) ,so organisational factors have a significant role in the empowerment of PHNs.

4.4 Social psychological; theory

Organisation structures do not empower (Laschinger 2001) but the employee's responses to the structures form the basis of psychological empowerment (Spreitzer 7995). Social psychological theory examines how to foster the process of individual growth and development. This theory supports empowerment by the removal of dis-empowering conditions. Authors like Kuokannen and Leino-Kilpi (2000) argue that this theory suits nursing as nursing is grounded entirely in human relations.

4.5 Critical social theory and social psychological theory

Critical social theory and social psychological theory both include reflection. Gilbert (1995) saw a problem with both in that they see power as located in the individual's minds and actions without identifying these individuals are products of power and how their identities are located within the material conditions of their lives and the social practices which operate there. Gilbert (1995) examined empowerment and the problem of power in relation to nursing and found that nurses need a clear conceptualisation of power and the ways it reduces conflict.

5. Literature review

5.1 Teaching critical social theories as an empowerment paradigm in nurse education

Fulton (1997) performed the first UK study to review nurses' views on the concept of empowerment and used a descriptive survey with focus groups and applied critical social theory and the work of Freire and Habermas to the enquiry. Their study group consisted of 16 nurses, some of them newly qualified from a variety of backgrounds. Four categories emerged from the focus groups: having personal power, relationship with the multidisciplinary team, empowerment, and feeling right about self.

The focus groups found nurses felt empowerment was a process and an outcome and conceptualised empowerment as freedom to make decisions with authority and to have choices and the process was reciprocal. The nurses did not feel they were empowered, and Fulton suggested that they showed signs of being oppressed. The nurses felt they lacked autonomy related to unequal relationships with medical staff. Their study recommended the teaching of critical social theories as an empowering paradigm in nurse education to promote reflection with action. Fulton claims this would support liberation and empowerment. Fulton's findings were in agreement with previous work by Gibson (1991) and Salvage (1992). The Health Service and the Educational institutions need to create environments for PHNs that create the conditions that make it possible for PHNs to reflect and act on their reality (Falk-Rafael 2004).

5.2 Empowerment by fostering ability

Haugh and Laschinger (1996) performed a comparative study using Kanters structural theory on the perceptions of power of two levels of nurses working in three public health agencies during a time of programme transition. A convenience sample of 46 PHNs and 10 nurse managers from three health centres in Ontario were surveyed as to their perceptions of the amount of power that existed in their own and their manager's jobs. PHN empowerment was found to be significantly related to their perceptions of their immediate manager's power in the organisation. Access to support was the factor perceived to be most lacking by nurse managers and Haugh and Laschinger (1996) identified that networking and seeking a mentor were substitutes they used. They distinguished in their research that managers perceived themselves to be more empowered in their work environments than the

PHNs had perceived themselves to be. The PHNs in their study presented a view that they perceived they had less access to empowering structures than their managers perceived they had. They (1996) suggested that nurse managers need to initiate efforts to improve access to power, opportunities, information, and resources for the PHNs. *"The study"* tallies with Kanter (1977) that in current organisational structures access to opportunities and power structures improve as one moves up the hierarchy. Therefore, PHN Managers need to initiate and sustain efforts to improve PHNs access to power structures and the Health Service needs to survey PHNs in relation to their access to power and opportunity structures and then act to remedy this.

Haugh and Laschinger (1996) argue that work environments must be structured in ways that empower PHNs to work effectively regardless of whether the focus is on individuals, families, or communities. So creating environments where PHNs are supported by empowering strategies will also contribute to the nurses and midwives working to their full potential and being empowered. Laschinger and Wong (1999) go on to argue that creating work environments that foster professional nursing practice by empowering nurses requires nurses to be accountable for their practice therefore supporting high quality patient care. Therefore, it would seem that there are political and professional, organisational and moral pressures on PHNs to work to support empowerment of their clients. In addition, if nurses are empowered in their work, the quality of patient care and the efficiency of the health services are improved. However, there are barriers, which constrain against the empowerment of the PHN and their role to facilitate the empowerment of their client.

McDermott et al (1996) used a descriptive correlation design survey applying Kanter's structural theory to examine the relationship between nurse's perceptions of job-related empowerment and their commitment to their organisation. Questionnaires were administered to a sample of 112 registered staff nurses in a 450 bed acute teaching hospital in Ontario. McDermott et al's (1996) study strongly suggested that nurses who perceive themselves as having access to resources, information, opportunity, and support in their work environments are more likely to be committed to the organisations. McDermott at al (1996) suggested that a vision of the nurse, as a knowledge worker is required to empower nurses and give them control over their own practice. All levels of nursing management must be genuinely committed to fostering empowering behaviours in staff, and this vision must begin with nurse managers (McDermott et al l996). PHNs may take the view that they have little power and this supports restricting their own professional freedom and their ability to support their client's power. When PHNs acknowledge the true sources of power as that generated through their actions, this will support a shift from the negative view of power. They will then be better positioned to take responsibility for their power and act on this power. To foster empowerment, Kanter suggests that power could mean efficacy and goal orientation and this he referred to as legitimisation. Robbins (1985) describes successful leadership with similar qualities. However, greater empowerment will be fostered if the Health Services also makes efforts to raise the self-efficacy of the nurses.

Barden et al (2011) conducted a descriptive correlational study to determine the relationship between perceptions of governance and empowerment among nurses working in an acute hospital setting in New York using Kanter's structural theory. These nurses were working in an environment where shared governance had been in place with six to twelve months. Their study found that their was a significant relationship between perceptions of shared governance and empowerment and this suggested that as shared governance progressed so did empowerment. The nurses rated the access to opportunity as the highest empowerment sub-scale, followed by information, support and then resources. Barden et al (2011) suggested that shared governance provides a vital communication and decision making infrastructure.

So, it would appear that the tools that create the environment for empowerment also support the maintenance of power by all, and the environment that is empowering is also one in which successful leadership is demonstrated. Nurse's empowerment is fostered by shared governance and open access to structures of power.

5.3 Acting on responsibility

Laschinger and Wong (1999) performed a cross-sectional correlational survey on a random sample of 672 registered nurses from a medical centre in Ontario. Their study was based on Kanter's model and linked Staff Nurse perception of workplace empowerment and accountability to two indicators of work effectiveness. Their findings supported the proposed impact of nurse empowerment on nurse accountability and perceived work effectiveness. It highlighted the importance of creating environments, which provide access to structures that empower nurses to be accountable for their work and to assume accountability for their client outcomes. The findings suggested that in order to be effective organisational leaders must ensure that nurses have information, support, and necessary resources to act with accountability. Also they suggested that nurses need true governance so that they not only need to be involved in the decision making process through committees and on, but they also need to have control over their professional practice to have trust in the organisation to support them acting with accountability.

5.4 Empowerment by raising self-efficacy

Conger and Kanungo, (1988) propose that to facilitate empowerment, strengthen the self-efficacy belief or weaken the belief in personal powerlessness. If, empowerment is based on Bandura's self-efficacy, the individual will have an internal drive to achieve and management focusing on strengthening self-efficacy belief can foster this. PHN will then be motivated and well positioned to self-empower. Kuokkanen and Leine-Kipli (2001) performed a qualitative study into the qualities of an empowered nurse and the factors that contribute to the empowering process. Their study used the social psychological theory in interviews with 30 nurses participating in a career advancement project in a University hospital in Finland. Their study suggested that power is generated through the individual's behaviours, actions, and relations to another person. The factors they found that promote empowerment are: confidence and collaborative skills, support and leadership The factors that prevent empowerment are: lack of moral principles and factors, different views on the goals of care and an authoritarian leadership style as it prevents participation and there is a lack of trust and sharing from management. Their study suggested that empowerment is a process steered by personal values and endeavours and environmental factors and an empowered nurse has those qualities which make possible high self-esteem and successful professional performance and progress. Notwithstanding, organisational theory has a significant impact on the empowerment of nurses, however an approach is needed that combines organisational and social psychological theory.

Kuokkanen et al (2003) used the social psychological theory of empowerment to perform a quantitative study into nursing empowerment, to identify the background factors associated with its realisation. "*The study*" group consisted of 200 critical care nurses from a university hospital, 200 long-term care nurses from 7 community hospitals and 20 trained PHNs from 25 health centres all health-care facilities were located in southern Finland. A questionnaire was developed, based on the moral principles and factors that were found to empower the nurse in their (2001) study. Their study found that empowerment had a significantly positive effect on affective commitment and trust in management. It also found that organisational commitment is highly related to job satisfaction. "*The study*" recommended that for empowerment to occur low hierarchical organisations were needed together with working practices orientated towards teamwork, coherent values and personnel management that creates opportunities and sufficient resources which acknowledges the influence of organisational theory on empowering work environments. Their study suggested that nurses need real influence and decision-making powers over the issues concerning their work.

Conger and Kanungo (1988) recommend the identification of conditions within organisations that foster a sense of powerlessness amongst subservients and their removal by empowering strategies and tactics (Conger and Kanungo 1988, p 474). Therefore, to raise the self-efficacy belief of employees raise their convictions of their own effectiveness. They also recommend personal efficacy information as vital to support this process. Empowering management practices involve the sharing of power and to be effective at individual level, must be perceived by employees as increasing their sense of self-efficacy, something they suggest a manager can accomplish through practices that are more informal, like verbal feedback (Conger and Kanungo, 1988).

Kuokannen and Leipo-Kilpi (2000) explored the empowerment concept to establish its compatibility as a theoretical framework for research into nurses' professional growth and development. They found that social psychological theory places emphasis on the individual and the environmental factors. Although, fostering the self-efficacy of the PHN in isolation may lead as Foucault and Freire suggest to a status quo. Self-efficacy needs to be fostering on an ongoing basis and needs to be supported by fostering ability.

5.5 Empowerment by raising self efficacy and fostering ability

Lewis and Urmston (2000) examined the concept of empowerment based on Kanters work (1993) and work on empowerment and the role of nursing and power relationships within organisations. They found that structural changes are not sufficient to allow greater empowerment greater effort is needed by all. Lewis and Urmston (2000) reflected on Seedhouse (1986) who viewed health as the ability to achieve human potential. They linked control over health and health promotion to organisational development and suggested that it was concerned with the provision of healthcare environments, which actively enable and encourage individuals to develop personal skills like internal locus of control, self-efficacy, and the ability to adapt. These environments foster freedom to practice with professional autonomy without too rigid authoritarian structures and will then support the fostering of the nurses self-efficacy.

Nurses need to be given the opportunity to participate more fully in decision making and to develop greater work ownership and power relationships where there is shared responsibility (Lewis and Urmston 2000). This calls for shared governance. Therefore, PHNs need their autonomy fostered and access to power and opportunity structure in their workplace to empower them.

5.6 Empowering environments and fostering self-efficacy of nurses

Laschinger et al (1999) reported that nurses perceived more structural empowerment and higher job satisfaction in their work settings when their leaders encouraged autonomy, expressed confidence in the nurse competence and facilitated participative decision making. Indeed Sabiston and Laschinger (1995) reported a statistically significant correlation between structural empowerment and work self--determination, and Wilson and Laschinger (1994) found a statistically significant correlation between structural empowerment and organisational commitment

Sun Ning et al (2009) conducted research to test Kanter's organisational empowerment model to examine the impact of nurse empowerment on job satisfaction using a cross-sectional design on a sample of 650 full-time nurses employed in six Chinese hospitals. They found a statistically significant positive correlation between empowerment and job satisfaction and concluded that nurses who view their environments as empowering are more likely to provide high quality care.

Laschinger et al (2001) performed a predictive non-experimental study using a questionnaire on a random sample of 412 nurses in Canada. Their study' purpose was to test Kanter's structural theory of power model and see if there was a link between staff nurse workplace empowerment, organisational trust, and perceptions of job satisfaction and organisational commitment. The model developed included measures of psychological empowerment. Trust they suggest comes from mutual respect, understanding and shared values. Whitney (1994 p 235) found that without trust individuals would not work together except under conditions of stringent control. Laschinger et al (2001) found a strong relationship between trust in management and nurses perceived access to information and support. Laschinger et al (2001) suggest that access to empowering structures was the strongest contributor to empowerment. They claimed that to be empowered nurses need real influence and decision-making power over work related issues, the work environment, and future projects. In addition, their research suggested that managers will need to focus less on control and more on co-ordination, integration and facilitation of nurse's work, together in a climate of trust. This fosters work satisfaction and genuine commitment to organisational goals to provide high quality client care.

Laschinger (2001) performed a predictive non- experimental design study including measures of psychological empowerment to test a model specifying relationships among structural and psychological empowerment, job strain and job satisfaction. The sample was a random sample of 404 Canadian staff nurses. They found that nurses need empowering organisational structures.

Underwood et al (2009) conducted research in Canada by focus groups with community nurses to identify what helps them and hinders them to work effectively and to identify organisational attributes that support public health nurses to practice to the full scope of their competencies. To practice effectively, community nurses need professional confidence, good team relationships, supportive workplaces and community support. To succeed PHNs need a combination of factors including sound government policy, supportive organisational culture and good management practices. Interestingly, management practices were identified as the most important attribute for effective public health nursing. Organisational attributes identified as supports for optimal practice include: flexibility in funding, program design and job descriptions. Management understanding and promotion of the role of PHNs, clear organisational vision driven by shared values and community needs, co-ordinated public health planning across jurisdictions, and strong leadership that openly promotes public health, values their staff's work, allowing room for creativity and invests in education and training and nurses to take advantage of learning opportunities. To support sustained competencies and confidence in their professional abilities, PHNs need more access to continuing education, policies, evidence and debriefing sessions.

Chang (2008) conducted a study in China to investigate the relationships between employee empowerment, innovative behaviors and job productivity of PHNs by cross-sectional research that examined for organisational empowerment based on Kanter's four empowerment structures and psychological empowerment based on Spreitzer's 12—item scale, for job productivity based on Hackman and Oldham 1975 and McNeeseSmith 1996 and Liou 2000 and for innovative behaviour based on Yeh (2000) and Scott and Bruce (1994). Their study found showed that psychological and organisational empowerment had similar predictability on productivity which was not in line with previous research which showed organisational to have greatest effect. The authors suggest that this may be cultural. In predictors for job productivity, the self-efficacy subscale in psychological empowerment showed greatest contribution. This was in line with previous research. However the negative relationship between the meaning of the job and job productivity were contrary to findings by Laschinger (2001) and Spreitzer (1996). The findings indicate that access to information in an organisation had positive effects on job productivity which was shown in previous studies (Laschinger and Wong, 1999).

Meagher-Stewart et al, 2010 conducted research in Canada to gain an insight into the organisational attributes that support PHNs to work effectively using appreciative enquiry through focus groups with PHNs, their managers and policy makers. The attributes they identified were government and other system attributes in particular sufficient and flexible funding and public health leadership and public health planning and co-ordination, local organisational culture, values and leadership characteristics including a clear vision, mission and goals for public health and these are understood throughout the organisation, fostering a culture of innovation and effective leadership that demonstrates respect, trust and support for PHNs and frontline management practices to include effective planning based on evidence and outcome focused, that the roles and responsibilities in public health be defined in relation to the overall goals and accountabilities rather than tasks, that PHN managers need to acknowledge PHN contributions and to build stakeholder understanding of the PHN role to all, supporting autonomous practice of PHNs, investing in ongoing professional development, training and education of PHNs and fostering this responsibility in PHNs, effective human resource planning, supporting community health partnerships including the time needed to build them, fostering effective communication and information sharing between PHNs and all levels.

Meagher –Stewart et al 2010 recommended the need for strong leadership throughout the public health system to support empowered public health nursing practice and in particular how their managers do their job is a key driver for optimising public health nursing effectiveness.

Therefore, PHNs need continuous efforts to foster their self-efficacy. They need to be supported to act on their expertise including more access to continuing education and support in order to have a sense of trust and this will support their perception of being respected and their self-empowerment. PHNs roles need to be understood and communicated effectively by their managers to other service.

5.7 Raising political consciousness, raising self-efficacy and creating empowering organisational structures

The Department of Health and Children (DoHC) commissioned a national study into the understanding and experience of empowerment by nurses and midwives in the Irish health services in 2003. The theory base for "*the study*" combined organisational, psychological and critical social theory approaches to empowerment. Their study consisted of two phases. Phase one comprised of focus groups held to elicit nurses and midwives understanding and experience of empowerment. Phase two involved a national survey of nurses and midwives understanding and experience of empowerment to test the findings of phases one. The factors that were found to enhance empowerment and these factors were also individual factors were; education, skills, knowledge and self-confidence (DoHC 2003, p 11). The factors that inhibited empowerment were environmental factors, which were poor management style, lack of education, lack of support, lack of recognition (from management and other professionals) (DoHC 2003, p 11). They discerned that the prerequisite for an empowering organization is individual factor conditions such as self-esteem and self-confidence, which can drive the formal and informal structures**.** The findings from their literature review, and the research study supported each other and provide a model for the empowerment of nurses and midwives that takes a holistic approach and includes individual, organisational and historical factors.

6. Discussion

The findings of "*the study*" of what participants perceive empowers and disempowers PHNs were in agreement with research into the understanding and experience of empowerment of nurses and midwives in Ireland by the DoHC (2003). The participants distinguished that PHNs need ongoing access to knowledge, education, training, and up-skilling to raise their self-efficacy to foster their empowerment. The DoHC (2003) also identified similar findings for nurses and midwives. The participants recognised open access to these factors as a sign that PHNs were supported and this support was equated by the PHNs to them being respected and trusted by the health service, their clients, and other healthcare professionals. The lack of ongoing open access to education, training, and upskilling was pinpointed by the participants as disempowering to PHNs. The DoHC (2003) also discerned that the lack of open access to education, training, and upskilling education was disempowering to nurses and midwives. Factors related to the organisational structures of power were distinguished in the research as of greatest impact on the disempowerment of the PHNs.

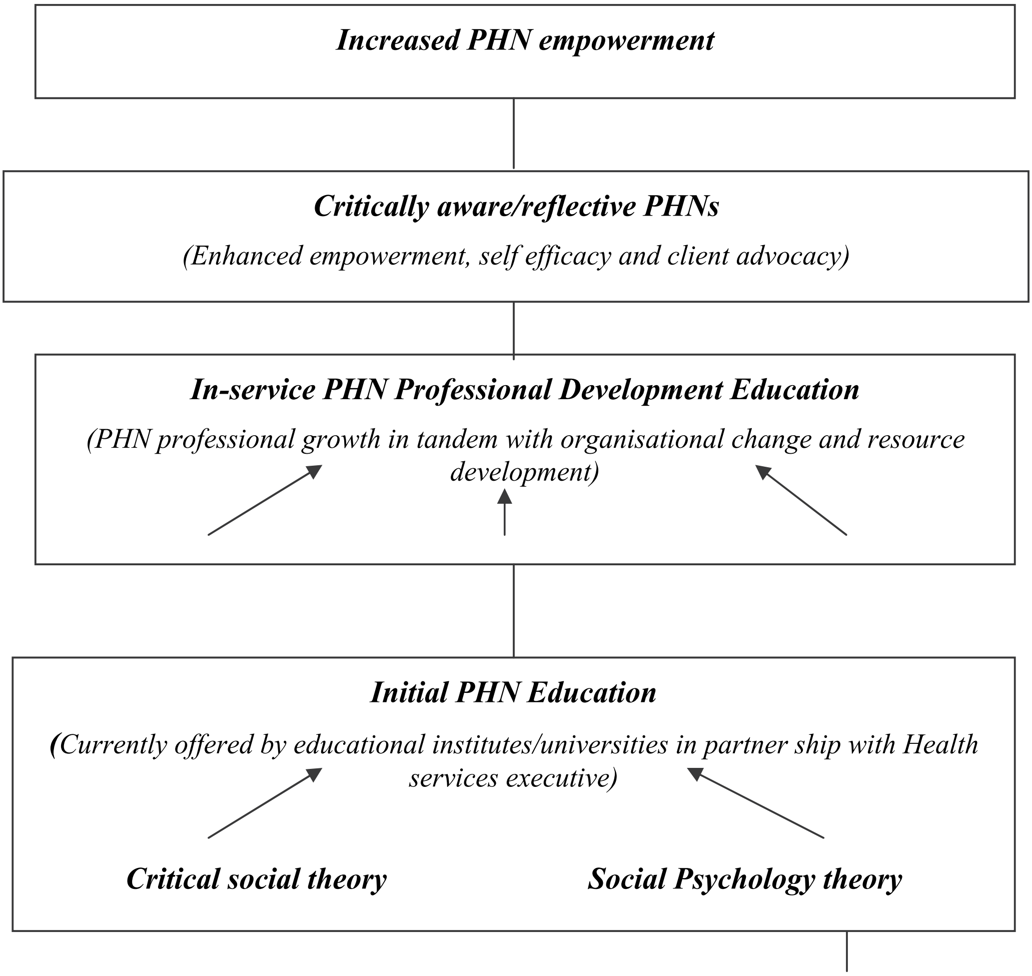


Table 1. The theory base proposed for empowerment of PHNs from *"the study"* (adapted from Cawley and Mannix McNamara, 2011)

Table 1 shows the theoretical base proposed by *"the study"* to support PHN empowerment. A critical social theory methodological approach is needed in the education, ongoing training, and workplace of PHNs. This approach must be coupled with organisational change specific to supportive resources and effective service provision by the health service. This approach is achieved by providing an organisational and social psychological theory approach to PHN empowerment. The potential outcome is the growth and development of the PHN. The process of facilitating client empowerment produces reciprocal empowerment including growth and development of the PHN.

6.1 PHN - role, workload and management structure

The traditional PHN role in Ireland was heavily influenced by a medical model approach to health with the PHN acting as an expert and telling the client what to do. This role was curative and preventative. Her responsibility included individuals of all ages requiring a domiciliary clinical nursing service, infants and children including school aged children, children at risk, individuals with physical and mental disabilities, young chronically sick individuals, those discharged from the psychiatric services and older individuals (Leahy-Warren 1998, p 14).

The current role of the PHN is unchanged in terms of population groups and is described as a generalist role encompassing, primary, secondary and tertiary care at the level of the individual, family and community. The traditional focus had been on individuals and illness however, the current focus is on population health and health promotion. The mission of Public Health Nursing (DoH 1997) calls for a health promotion approach to practice and calls for PHNs to work in partnership with their clients. Despite this, the current job description in Ireland as outlined in Circular No 41/2000 (DoHC, 2000) confines Public Health Nursing primarily to a curative role in a series of tasks.

Zerweck (1992) suggests that the central purpose of PHN home visits is to encourage individuals to self-help and to foster their responsibility. Therefore, developing relationships with the community is fundamental to the role and this requires the resource of time. So, the health promotion approach needs the resource of time and interpersonal and communication skills. The needs for PHNs with regard to time and time management need to be prioritised by Health Service. Interpersonal and communication skills need to be fostered in the education and upskilling of the PHN and PHNs need to be facilitated through continuous development processes to include opportunities for their personal development to support this.

Despite larger populations, clients with more complex problems and earlier discharge from hospitals the geographical areas have remained unchanged in Ireland. This method of workload distribution has been heavily criticized in the literature (Hanafin et al 2002), because it masks different population demographics, different population needs and locally available services and social support (DoHC 1997). Chavasse (1998, p 174) criticized the large caseloads of PHNs as hampering their ability to 'provide primary as well as secondary nursing care.' Workload was identified by McDonald (1993) as an impediment to PHNs work to support client participation. Hanafin (1997) reiterated her concern by suggesting that the large caseloads of PHNs were eroding their management and health promotion role. Added to this, a significant weakness of Public Health Nursing is that 'the curative dimension of their work took precedence over the preventative' (O Sullivan 2005, p 38).

PHNs have a large clinical role and the literature suggests that as the clinical role demands increase, the health promotion role diminishes. Discursive analysis of the focus groups in *"the study"* identified language suggestive of a lack of freedom related to the demands of the role and the PHNs voiced that this showed a lack of respect to them. Participants also evidenced language related to

The PHN workload may interfere with their power at an individual level resulting in weakened work autonomy. The impact identified by the focus groups in *"the study"* was PHNs feeling that they were overstretched, disrespected and disempowered. The lack of taking responsibility for advocating for their role and for not becoming more political in their role and not acting accountably may be related to this disempowment. Although, PHNs may be unaware that their autonomy is affected. This lack of autonomy therefore may contribute to PHNs reverting to their traditional expert role to restore their autonomy.

PHNs need to be facilitated by Health Service to reflect on their role and the philosophy behind the role to realize the role calls for and needs a health promotion approach. In addition, the continued professional development of PHNs needs to include leadership skills to support PHNs to lead an empowering practice and power modules to support their critical reflection on how power operates to foster empowerment.

The service provided by the PHN is not only influenced by the needs of the community, but also by the level of other services (both nursing and non-nursing) provided in the community (Hanafin et al 2002, Begley et al 2004). By filling the gaps, PHNs support the masking of health service gaps and inequities and this reduces time for their health promotion work. It is imperative that the role is clarified. The range of activities and responsibilities of the PHN is broad and can lead to difficulty in articulating the boundaries of the PHN role (Begley et al 2004). In the report "An agenda for the future development of Public Health Nursing", (DoHC 2005) the need for role clarity and PHN concern about the demands of their workload (DoHC 2005, p 13). Many studies acknowledge that the workload and the organisation of the PHN service place unreasonable demands on the PHN (DoHC 2005, Begley et al 2005). These studies together with repeated attempts to define the role of the PHN suggest that the role is contested and that a reorganisation of the PHN service is required. The consequence of an unclear role and a large workload is that arguments can be made for not meeting the accountability of the role.

Fostering a health promotion approach to practice requires change management. Backward mapping is a bottom up approach that begins at the point at which administrative actions intersect with private choices and it emphasises that it is not the policy or the policymaker that solve the problem but the person closest to where the change is needed that does (Elmore 1989). The crucial difference between forward and backward mapping is that the forward mapping process relies primarily on formal devices of command and control that centralise authority and backward mapping relies primarily on informal devices of delegation and discretion that disperse authority (Elmore 1989, p 249). Backward mapping is the most effective way of ensuring a relevant and successful change implementation process (Elmore 1989). Elmore argues strongly that those nearest to the point of change have the greatest ability to influence it and the problem solving ability of complex systems depends not on hierarchical control but on maximising discretion at the point where the problem is most immediate (Elmore 1989, p 248). Therefore, those nearest to the point of change are best positioned to inform change agents of what changes are necessary and also of the most successful strategies for change (Ryan et al, 2006).

Because of PHNs workloads, unclear role together with innovation overload related to the expanse and expansion of the role, PHN could see themselves as the only practitioners delivering services in primary care. Innovation overload is a term used commonly in education fields related to an obsession with understanding or creating what makes consumers happy (Business Innovator Insider, 2006). The outcome is that PHNs have a huge body of expertise to meet their ever-expanding role and therefore do not want to let any part of their role go, as was evidenced in *the study*. Unfortunately, the consequence of holding on and not letting go of any part of the role is that the health promotion potential becomes limited. Interestingly, the literature evidenced that PHNs had mixed reactions to the entry into the community of RGNs. The 'Commission on Nursing' (DoHC1998) recommended a permanent role for RGNs in the community care team in Ireland, in line with public health nursing service needs and circular no 41/2000 (DoHC 2000) asks PHNs to share power. RGNs can act as a threat and a resource because they perform tasks that have previously been performed by PHNs and PHNs have invested heavily in the role and the skills to support the role.

A revised strategy to ensure a sustainable community nursing and midwifery service that will effectively meet the health needs of the population of Ireland within a primary care setting (DoHC 2001) is awaited. This report should give a vision for the role of the PHN and he report should give a vision for the role of the PHN. Added to this are the moral, political and professional pressures and barriers on public health nurses working to the health promotion model. The impact of the many and varied influences on the role, its development and the workload of the PHN is that there is a lack of an integrated vision for the role of the PHN. The outcome is that PHNs feels undervalued, misunderstood, disempowered and responsible for all the needs in primary care and this makes the health promotion role impossible to fulfil. This is related to the disempowerment of PHNs and then PHNs may revert to their traditional expert role, which interferes with empowerment of the client. Therefore, it is fundamental that PHNs have a clear vision for their role and that their roles are clearly defined with equitable caseloads. Nic Philibin et al (2010) examined the role of the public health nurse in one Irish community care area to try to clarify the role of the PHN and they found PHNs have a role within each facet of community health. They recommended that PHNs need to define and redesign their role so that they no longer think that they are the catch-all in the community and so they can respond to the changes in need of their clients.

6.2 Health service management structure

The Health Service in Ireland has a characteristically hierarchical and bureaucratic management structure with attempts to flatten the structure as part of the current restructuring process. Bureaucratic systems are characterised by authority and accountability attached to roles with authority passing from senior to junior and accountability going from junior to senior (Bush 1995, p 43). This system is often associated with delegated power without acknowledging individual contributions. The management structure for Public Health Nurses is positioned inside this structure. PHNs are expected to work as autonomous practitioners with line management from their Assistant Directors of Public Health Nursing (ADPHN) and accountable for their practice to the Director of Public Health Nursing. Assistant Directors also have a role to support the PHNs.

Argyris (1964) was one of the strongest critics of bureaucracies and he argued they restrict the psychological growth of the individual and cause feelings of failure, frustration, and conflict. He (1964) suggested that organisational environment should provide a significant degree of individual responsibility and self-control, commitment to the goals of the organisation, productiveness and work, and an opportunity for individuals to apply their full abilities. This approach could foster the autonomy of the healthcare professional. It also add to calls for shared governance in the practice of Public Health Nursing together with the regular review of management systems to ensure they foster the growth and development of PHNs and meet the needs of the clients of the health service. Governance is not only the need to be involved in the decision making process through committees and so on, but also the need to have control over one's own professional practice. PHNs work as autonomous practitioners and yet they are so entrenched in their ever-expanding role that their opportunity to develop their autonomy is hampered. In addition, the lack of vision for the development of the PHN role (including the recognition of the role of the PHN as specialist and the development of PHN practitioner roles) interfere with the development of governance in public health nursing. Notwithstanding this, application has been made for the role to be recognised as a specialist role, and this has to date not been approved. Clarification of the role of the PHN should support the recognition of the role of PHN as specialist and create a foundation for the fostering of governance in public health nursing. Ongoing and regular review of management systems is a normal part of management and needs to take place within the entire structure of the Health Service with the focus on the needs of clients of the health service. The participants in *"the study"* perceived that they needed their organisation to facilitate their empowerment. Clarke and Mass (1998) examined how empowerment and collaboration when actualised impacted on clients and nurses in British Columbia and recommended that if nurses are to integrate the principles of empowerment and collaboration in their practice, changes in healthcare reform must focus on vesting professional nursing organisations with sole authority to govern nursing, granting nurses authority for aspects of care necessary for care within the scope of advanced nursing practice and revising of reimbursement mechanisms to pay for these essential health care services.

The participants in *"the study"* identified that the Health Service did not empower them through political constraints on the services they were trying to assist their clients to receive. This suggests that PHNs feel dis-enfranchised themselves by the health system. Haugh and Laschinger (1996) used the organisational theory of power to examine empowerment of two levels of Public Health Nurses and found that nurse managers had overestimated the level of empowerment and the access to empowering structures of the PHNs. They suggested that managers need to initiate efforts to improve access to empowering structures for PHNs.

McDermott et al (1996) used a descriptive correlation design survey applying Kanter's theory of power to examine the relationship between nurse's perceptions of job-related empowerment and their commitment to their organisation. They found that if nurses perceive they have access to empowering structures at work they are more committed to the organisation. They suggested that self-efficacy can be fostered by providing feedback on performance, allowing nurses to have autonomy over their work and developing mechanisms to raise the profile of nurse's work as knowledge workers.

The participants in '*this study*' affirmed that the self-efficacy of the PHN needs to be maintained and affirmed and giving PHNs ongoing access to knowledge, training, and upskilling fosters self-efficacy. The participants acknowledged that this was recognised by them as a sign of respect and support for their role from the Health Service, their managers, clients and other healthcare professionals. They related this support as a recognition of the value of their role and this support was both formal by organisational structures and informally by feedback and affirmation by all, on their role. So, it is fundamental for the Health Services to examine the access PHNs have to empowering structures and make efforts to support improving access, feedback, support and self-efficacy of PHNs.

Laschinger and Wong (1999) performed research into nurse empowerment, organizational theory approaches to power and accountability and found that they need to open access to structures of power, fostered empowerment, accountability, responsibility and effectiveness. They (1999) suggested that to support this, managers at all levels need to relinquish traditional controlling roles. The focus groups in *"the study"* related this to having flexibility in their time management and autonomy to manage their roles. Laschinger and Wong (1999) also suggested that nurses need true shared governance. They (1999) also suggested that PHNs need to have trust in the organisation and this supports them acting with accountability**.** Laschinger et al (2001) and Laschinger (2001) examined the link between staff nurse workplace empowerment, organisational trust, and perceptions of job satisfaction and organisational commitment using Kanter's model and incorporating social psychological theory. They found that access to empowering structures was the greatest contributor to empowerment. This access was equated to managements' trust. The participants in *"the study"* equated trust with their access to knowledge training and flexibility in their job management. Therefore, access to empowering structures is needed to foster accountability. This suggests that the organisational theory approach to empowerment acts to raise the ability and the self-efficacy of the PHNs. The participants in *"the study"* related their sense of empowerment to their access to access to knowledge, education, training, and support.

6.3 Knowledge

Knowledge was identified in *"the study"* as one of the cornerstones to PHN empowerment and fundamental to PHN autonomy and self-esteem. The participants recognised knowledge as essential to PHNs credibility and to their empowering role. Participants identified that knowledge raised their self-efficacy in their role and without it; they could not facilitate the empowerment of their clients.

The PHN is often referred to as a knowledge worker, because of the body of knowledge and the wide expanse of the role. McDermott et al (1996) suggested that a vision of the nurse, as a knowledge worker is required to empower nurses and give them control over their own practice. Knowledge according to the participants in *"the study"* gives the PHN power. This power is expert power. Having expert power implies having the necessary knowledge, competence and skill to perform their role effectively within the context of caring (Benner 1984). Interestingly, there is an understanding from the literature as the PHNs role continues to enlarge her concomitant knowledge and related expert power is raised and there is no letting go of any parts of the role. Foucault examined power and knowledge and argued that they were closely interwoven (Foucault 1980). He argued that the reorganisation of knowledge created new forms of power and this he called disciplinary power. Ryan et al (2006) remind us that expert power has the potential to disempower clients by oppressing them, through clients not understanding medical knowledge and medical discourses. Foucault (1980) claimed that power is everywhere and that it is distributed throughout complex social networks so that individuals are the means and the generators of power whether they are a healthcare professional or a client of the health service. If the 'Ottawa Charter' tenet of social justice is applied, then there is no place for an expert power approach and what is preferred is the empowerment approach, which is the client centred approach.

However, PHN power is anchored in expert power related to their knowledge. Therefore, PHNs need to relinquish their expert power to support working in partnership. This task is becoming increasingly difficult as the role continuously enlarges and therefore the knowledge base explands. Alongside this, PHNs do not share parts of their role.

Relinquishing power is achieved by PHNs questioning the structures and processes of power and this will lead PHNs to challenge their expert power and how this can disempower their clients. This process can be fostered through education in critical social theory and power modules and the facilitation of critical reflection by PHNs in their workplaces. The knowledge base of the PHN is still needed as a resource to be used in the context of the client centred partnership approach. Therefore, the focus in health promotion is not on building capacity with regard to power over another but rather is concerned with power to engage in action (Carey 2000). The relinquishment of expert power holds the potential for the PHN of a new form of power called intellectual liberation.

The literature confirmed the ability of the PHN to see the bigger picture and suggested that this is based on their broad educational base and knowledge of the community resources. Intercommunity knowledge (Hanafin et al 2002) is the knowledge that PHNs gain in practice and is fundamental to their health promotion approach. The huge investment by the Health Service into knowledge to support the PHN may negate against the PHNs intercommunity knowledge. The danger is that the determinants that affect health may be ignored by disqualifying intercommunity knowledge. In addition, for PHNs their intercommunity knowledge becomes of less importance to them and public health nursing practice fails to develop its own theory base. PHNs and their managers need to reflect on both sources of knowledge to bridge the knowledge practice gap. O'Halloran (1998, p 113) suggests that the generation of knowledge to support and inform the practice of nursing must by its very nature emerge from that practice. Action research holds the potential to improve practice, generate new theory, and bridge the gap between knowledge and practice.

Cohen and Reutter (2007) reviewed scholarly literature from Canada, United Kingdom and United States of America to ascertain support for PHNs roles in reducing poverty and its effects and then reviewed professional standards and competencies in Canada to support the development of the role of PHNs in addressing child and family poverty. They concluded that if given more organisational support and enhanced appropriate knowledge and skills, values, beliefs and attitudes this PHNs role would be supported. Development of PHNs that includes critical social theory needs to be facilitated by their educational institutes and workplaces.

6.4 PHN managers role

For PHNs to develop their health promotion role to a greater extent the Commission in Ireland on Nursing (DoH 1998) contended that PHNs need to have their professional autonomy supported by the public health nursing management system. Notwithstanding this, Begley et al (2004) who conducted research into the role and workload of PHNs identified that the Assistant Directors of Public Health Nursing' (ADPHN) work is primarily management, rather than leadership. A leadership and a support role is inferred in the role of the ADPHN, but not mandated The participants in 'this study' were explicit in identifying the role of their Assistant Directors as to provide them with support in relation to issues that arose in practice and providing them with flexibility around their time to meet the job demands. Although, when the participants in "*this study*" discussed the support, they received at work, they referred to relying on work colleagues and taking time out. This suggests that despite the need for their Managers to provide them with support, in practice they look elsewhere for this support. Kuokkanen and Leino-Kilpi (2001) examined empowerment and found that a democratic leadership style was needed and when support is fostered this raises the empowerment of the nurse and the opposite is true. The democratic leadership style is one where the needs of the healthcare practitioner are the driving force. Interestingly, Directors also claimed they themselves had a lack of leadership at leadership level (DoH 1998). Leadership skills were identified as needed by all grades of PHN (Begley et al 2004).

Kuokannen et al (2003) examined empowerment using the social psychological theory and found that if nurses are empowered this correlates with them having trust and commitment in, and to the organisation, to their professional activity and job satisfaction. Mullins (1998, p 58) suggests that with rapid changes in the external environment, de-layering of the organisation, empowerment, and greater attention to meeting the needs of customers, there is an increasing need to organise for flexibility. PHNs also need to be flexible with their clients and therefore need flexibility from their managers. 'Flexibility and responsiveness will be essential as we continue to meet a progressing health service' (Armstrong 2000, p 138). The participants in *"the study"* identified flexibility in time and job management as key and this requires nurse management trust.

The participants in *"the study"* also identified that they needed support to raise their self-efficacy and training and education in isolation were inadequate. This support includes promoting the vision of the PHN as a knowledge worker, raising the profile of PHNs and ongoing feedback and support for the PHN. Do Assistant Directors have time to provide support to the PHNs and to work to raise the profile of the PHN? In relation to the continued professional development of PHNs, the Assistant Director had a vital role to play.

6.5 Self-efficacy

Self-efficacy is based on the idea that power lies within the person and education, training, and practice of PHN skills and support are essential to this. PHN self-efficacy needs to be maintained to foster their self-efficacy. The participants in *"the study"* equated their self - efficacy to their empowerment. The participants related their lack of self-efficacy to a lack of respect. The research suggests that respect can be fostered by enabling open access to the organisational structures of opportunity and power.

The discursive analysis of *"the study"* evidenced language suggestive of PHNs beginning to strive for liberation related to the PHNs perceiving and reflecting on their perception of feeling intimidated by other professions, and their perception of their lack of profile, in comparison with hospital nurses. The discursive analysis also pinpoints language suggestive of PHNs becoming liberated related to the PHNs feeling and reflecting on feeling misunderstood and undervalued and moving onto PHNs acknowledging their expertise. The research affirms that PHNs needed to acknowledge their expertise, advocate for, and ask others including their clients to advocate for their role to foster their self-efficacy.

Rose and Glass (2008) reviewed research on emancipatory research methodologies to highlight the importance contribution they can make to the ongoing development of contemporary nursing practice by exploring the relationship between the nurse's emotional well-being and their professional practice. They found that when nurses are involved in research that is focused on their own emotions and / or performance at work, such as 'from the heart' issues, this can result in states of empowerment and personal liberation and these processes can transform practice.

The self-esteem of the PHNs is an important contributor to her autonomy and her self-efficacy and needs to be fostered by the Health Services and the PHN herself. The self-esteem of a person can be fostered by for example, pastimes, healthy lifestyles, stress management and life skills development amongst others. Fostering healthy self-esteem supports, a healthy internal dialogue for the PHN that will help when faced by work and personal life demands that impact on her emotional wellbeing. Therefore, PHNs will need to be more open about their need for and lack of support and this process will need to be fostered by their managers.

Conger and Kanungo (1988) advocated the social psychological theory to foster empowerment and suggested that empowering management practices should involve the sharing of power. They suggested that to be effective at individual level, they must be perceived by employees as increasing their sense of self-efficacy, something they suggest that can be accomplished through practices that are more informal like verbal feedback. They also suggest that to foster empowerment strengthen the self-efficacy belief or weaken the belief in powerlessness. Chandler (1991) examined and demonstrated that emotional support was the most relevant antecedent to empowerment when the empowerment of 267 staff nurses was surveyed. Feeling supported and affirmation was identified as a factor in raising self-efficacy by the focus groups. This suggested that there was a lack of support from the Health Service for the PHN and this was identified as a recurrent theme in the research findings. The participants in *"the study"* identified that support is needed from the Health Service and public health nursing management and that this fosters PHN self-efficacy. Although it is worth bearing in mind that the participants were all female and this may have impacted on their need for support.

Notwithstanding this, if PHNs do not feel empowered they may unknowingly retain their expert power to maintain their autonomy. However, PHNs also need organisational structures to focus more on ongoing support from the health service, feedback on their performance and affirmation to foster their self-efficacy and motivation. Self-efficacy of the PHN can be further fostered by PHNs evaluating their work, and taking an evidence based approach to their practice. Therefore an organisational and a social psychological theory approach to empowerment are needed in the workplace that includes mechanisms to foster support for the PHN.

The central themes emulating from the discussion are related to how to support PHNs in the change process related to relinquishing power to work in client centred partnerships with their clients. The change process is predicated on the empowerment of the PHN, which is essentially restoring and fostering the autonomy and self-efficacy of the PHN on an ongoing basis**.** The central themes are clarification of PHN role and workload, open access to organisational factors, time and its management, continuous professional development of PHNs and raising the self-efficacy of PHNs. This will require the Health Service to foster the clarification of the role and workload of PHNs and to foster the ability and self-efficacy and empowerment of PHNs to meet their empowerment role.

6.6 Clarification of PHN Role and workload

The PHN role including her health promotion role and the vision for the future of the PHN role is unclear. The mandate for PHNs does not make clear how PHNs are to work in a health promoting way. To enable the empowerment of PHNs, the Health Service needs to actively involve PHNs in the decision making process, related to their role development (Clarke, 2004). Redefining the role will involve letting parts of the role go and sharing the role with other healthcare professionals. Perhaps a more specialist role (Nic Philibin et al 2010). This, together with workloads being distributed by a more equitable model will support the workload of PHNs becoming more manageable. This needs to be supported by open access to organisational structures of empowerment. The outcome will be increased self-efficacy and empowerment of PHNs.

6.7 Open access to organisational structures

Structures of power include open access to all the knowledge and information one needs to know to do the job, all the resources needed to carry out the job efficiently and support and feedback on performance and continuing approval of actions. Public Health Nurses needs to be surveyed as to their access to the organisational structures and to identify the exact resources (including support) that they need. The outcome of this open access is the empowerment of PHNs and the added bonus to the Health Service is their commitment, motivation, accountability, responsibility, and effectiveness are fostered.

Laschinger and Wong (1999) suggest that managers need to relinquish control and work to support co-ordination, integration, and facilitation. This supports PHNs working as autonomous practitioners. The Health Service needs to review its management systems and make efforts to support PHN managers to relinquish their control and support teamwork. Lewin and Urmston (2000) suggest that structural changes are not sufficient; nurses need to be supported with leadership skills.

Lewin and Urmston (2000) and Laschinger and Wong (1999) and Barden et al (2011) all recommend the development of shared governance to give nurses real autonomy over their practice and freedom to practice autonomously without management interference. This fosters ability and self-efficacy. Therefore, the Health Service needs to take steps towards addressing the need for leadership skills and shared governance for PHNs. This can be fostered by supporting the clarification of the PHN role and changing the workload distribution to a more equitable one. The Health Service also needs to reflect on management practice and critique itself on an ongoing basis to ensure it meets the growth and development needs of PHNs and the needs of the clients.

6.8 Continued professional development of PHNs

Continuous professional development is needed for PHNs to meet the demands of their role. This needs to include personal development to support PHNs in fostering and developing their client centred partnerships. Personal Development planning is a suitable mechanism to ensure the growth and development needs of the PHN are linked in with the needs of the Health Service. This will ensure PHNs develop the skills they need to meet the demands of their evolving role and their professional and personal growth plans. It also ensures the Health Service is developing its main resource to support the empowerment of clients.

Efforts at fostering and maintaining self-efficacy are fundamental to support the continued professional and personal development of the PHN. Modules are needed in health promotion and power and health promotion needs to be integrated into the role of the PHNs. PHNs need training in evidence based practice and critical reflection skills.

Traditionally PHNs had high levels of expert power related to their traditional role and their body of knowledge. Working to promote health calls for PHNs to partner their clients (Chavasse 1992). As PHNs are called to work to the health promotion model, their role remains unchanged but enlarging, therefore they receive more knowledge, which supports their expert power, and they do not share any part of their role. Added to this, their caseloads are becoming greater with more complex needs leading to an increased focus on the curative aspects of the role. Coincidentally the PHNs nursing mandate and national strategy in Ireland do not make it clear how the PHN is to promote health. The result is increased expert power, reduced time to work to promote health, and reduced PHN autonomy and subsequent disempowerment. The consequence is the PHN may revert to their expert role and disempower clients and reciprocally disempower themselves.

The research findings from "*the study*" made little reference to PHNs analysing power to foster their empowerment and to support PHNs to relinquish power to support their empowerment role. Ryles (1999), and others suggest that nurses need to relinquish power. This calls for a critical social theory approach in education and practice to support PHNs. Sharing their expertise and partnering the client can liberate the PHN from feeling responsible for their client's outcomes. A critical social theory approach to the education of PHNs and as a methodology in practice should foster PHNs acting on their expertise and taking responsibility for their power. By reflecting on power as it relates to their role, this will support a process of conscientization that should lead to action. This action is the PHN taking responsibility for her own power and will support raising the autonomy of the PHN. This will support PHNs to analyse power and will lead PHNs to reflection on how they as PHNs are constrained by power and how they may use power in ways that do not support empowerment.

Reflective practice is needed to help PHNs to continue to analyse power and to continue to meet the needs of their evolving role. Also, Clarke (2004) suggests that raising the power of PHNs can be supported by PHNs getting a seat at the table of power, by making their invisible work visible, by advocacy for the role of the PHN by all, by the political education of PHNs and by PHNs receiving training with other members of the multidisciplinary team.

Simultaneously, PHNs need to raise their voice by other means. The focus groups in "*the study*" argued that PHNs need to advocate for their expertise, accountability, professionalism, and workloads. PHNs need to take responsibility for this body of expertise and act on it in order to support PHNs as autonomous practitioners. Raising their profile will raise the power of PHNs. To raise the profile of PHNs, Clarke (2004) and Hanafin et al (2002) suggest that PHNs need to make visible their non-quantifiable work because otherwise policy makers and practice developers will conceptualise their roles into simplistic tasks and ignore the special, creative, and artistic aptitude, capacity and strength which is public health nursing. This will improve the recognition and status of the PHN by other professionals and clients (Clarke 2004). PHN Managers are ideally placed to foster this.

Foucault (1994) suggests that the disposal of power requires the assumption of some responsibility. PHNs have power by virtue of their knowledge, skills role, and location in relation to their clients. Illich (1970, p 18) suggests that, freedom and power are determined by a willingness to accept responsibility for the future. When the PHN uses her power, she is assuming responsibility. The current PHN mandate acknowledges PHNs autonomy and asks for PHNs to take and to share their responsibility. Notwithstanding this, PHNs are unlikely to advocate for themselves to support raising their power until they have a clear vision of their future. Having a vision for the role of the PHN will give the PHN confidence to reflect and challenge their current use of power and to be able to relinquish power.

6.9 Time as a resource and its management

Time was identified as a theme in *"the study"* preventing PHNs engaging in health promotion practices. The lack of role clarity has a huge impact on PHN time and these need to be clarified to ensure PHNs are clear on their role and its boundaries. The clarification of the role will lead to delegation of parts of the PHN role. PHNs need to be open to the needs for the role changing to meet the demands of the service including the health promotion approach. When the PHNs role is clear work, they can prioritise to facilitate a health promotion approach. The PHN workload is expansive and again needs to be reviewed to ensure equity in workload distribution.

The arguments regarding the lack of time are preventing the role and profile advancement of Public Health Nursing, and interfere with her self-efficacy. Time is a resource and time in motion studies will identify the time resource needed for the PHNs nursing tasks (including the time needed for the health promotion approach and non quantifiable work). Time needs to be planned and managed by PHNs. Claiming there is a lack of time is an argument for not engaging in health promotion but may be a symptom of resistance to change. To support this resistance, action sets are needed to facilitate reflection on the evolving role of the PHN and the National vision for the role. In addition, continued professional development of PHNs is needed to assist them in developing a healthier approach to power. The vision of the PHN is fundamental to fostering this approach to power.

6.10 Raising self-efficacy of PHNs

Some PHNs in *"the study"* expressed that they perceived themselves to be disempowered by their clients and by other healthcare professionals, but the scope of the research did not allow the researcher to evidence this. Efforts at raising the self-efficacy of the PHN will support this. Efficacy raising can be fostered by PHNs being supported through a combination of raising ability by the Health Service and efforts at raising self-efficacy by their managers. This requires their managers providing them with ongoing feedback on their performance and the development of mechanisms to raise the self-efficacy of PHNs on an ongoing basis like acknowledging good projects that PHNs develop and giving PHNs an opportunity to present their work to their colleagues and others. This will also foster the self-esteem of the PHN related to supporting her emotional needs. A democratic leadership style is also advocated to foster the self-efficacy of PHNs (Kuokkanen and Leino-Kilpi, 2001). PHNs need real influence and decision-making powers in their work and this calls for shared governance (Kuoknnen et al, 2003). This needs to be supported by ongoing professional development of PHNs.

7. Conclusion

Nightingale 1891 looked forward to the time when nursing would be no more for the sick but for the well. Empowerment is the approach proposed to promote health by efforts at restoring the autonomy of the individual. The empowerment of the PHN affects her autonomy and this affects her engagement in empowering practices. Fostering empowerment is related to supporting PHNs to relinquish their power and develop new power that support their growth and development and that fosters empowerment of their clients. The expert power of the PHN is related to their ever increasing knowledge base and their disempowerment is related to the expanse of the role and workload. The outcome of not relinquishing their expert power is that PHNs retain expert power and this interferes with the process of client empowerment.

To support empowerment of the PHN a multipronged approach is needed. The role and the workload of PHNs need to be reviewed and updated to reflect the challenges posed by the health promotion role and the reality of the barriers to this role. This should support the resource of time to health promotion. This should support raising the self-efficacy of the PHN but support from PHN managers is fundamental to sustain and maintain the PHNs self-efficacy. The management systems in place in the Health Service need to be analysed on an ongoing basis to foster management systems supportive of empowerment.

The process to support PHNs relinquishing power requires change management and Elmore's (1989) backward mapping process is proposed. This requires a critical social theory approach in their education and practice, to support their critical reflection on power and their role. This will support PHNs to understand power and support them to foster their own and their client's empowerment. Together with modules on power. PHN managers need to work to raise the PHN profile and to support the role becoming more visible. PHN managers need to foster the self-efficacy of PHNs on an ongoing basis by supporting PHNs to practice autonomously, acknowledging their work and working collaboratively with them. Shared Governance provides a vital communication and decision making infrastructure (Barden etc al 2011) in empowering PHNs.

The health promotion role of the PHN will require the continuing professional development of PHNs. This will need to include their personal development to support their maintenance of their self-esteem and to support the client centred approach.

Support and leadership were identified, as needed by PHNs and the research raised concerns over where this support and leadership is to come from. The support also needs to include mechanisms to foster the self-esteem of the PHN. How can the Health Service support raising the self-esteem of PHNs? Currently the support and leadership role is not allocated and defined. This is a gap to the empowerment and the empowerment role of PHNs. The research identified the ADPHN as one of the key people to provide this leadership and support.

The main issue emulating from the research is that PHNs need organisational empowering structures including ongoing support together with ongoing efforts to raise their self-efficacy (Cawley and Mannix cNamara 2011). This process is a change management process and requires the relinquishment of power by PHNs, their managers and the Health Service. By becoming aware and remaining cognisant of the fundamentality of the need for the client to be at the centre of care the health promotion approach becomes the only way forward.

7.1 Study recommendations

"The study" recommends that PHNs needs to be involved in a process to review their role and workloads and to identify their resource and support needs to enable them to work to promote health. Continuous professional development is needed for PHNs including the personal development of PHNs. This needs to be part of structured personal development plans that feed back into the Health Service to ensure the Health Service develops appropriate training to meet the needs for the health promotion approach. The self-efficacy of the PHN needs to be fostered on an ongoing basis. (Recommendations are in Appendix 1) A suggestion for further research arising from this research is that the Health Service survey PHNs as to their access to empowering structures.

8. Appendix

8.1 Appendix 1

Recommendations for PHN nurse managers

PHN role and workload

- PHN Managers need to engage in consultation with PHNs in order to gain clarity with regard to exactly how problematic workload issues are.
- PHN managers need to lobby for PHNs with regard to a comprehensive analysis of the PHN role and workload and solution to these issues.

If PHN roles and workloads can be altered, then PHNs will have time to engage in health promotion with their clients and this will decrease clients drawing on PHN time in the long run.

PHN access to empowering structures (including management structure)

- PHN managers need to initiate efforts to improve PHNs access to empowering structures.
- The management structure and management styles need to be reviewed on an ongoing basis to meet the growth and development needs of PHNs in order that they meet the needs of their clients.
- PHN managers need to foster democratic leadership styles.
- The debate about shared governance needs to be prioritised to foster the autonomy of PHNs.

PHNs equate access to organisational structures to having trust in management. Improved access, fosters PHN empowerment and is related to PHNs professional activities, PHNs having greater commitment and job satisfaction.

Continued professional development of PHNs

- PHN Managers need to prioritise professional development for PHNs in the area of empowerment, time management, and evidence based practice.
- PHN Managers need to encourage and facilitate PHNs to maintain personal development plan that feed back into the Health Service.
- PHN Managers need to facilitate PHNs engagement in individual and group reflective practice sets to identify their training needs in relation health promotion.

- PHN Managers need to include personal development in the continued professional development of PHNs to equip PHNs to develop their skills in client centred partnerships.
- PHN Managers need to support multi-disciplinary training for PHNs.

Managers need to demonstrate their commitment to PHN empowerment by supporting professional development for PHNs and supporting PHNs by giving them time to develop and critically analyse the implications of empowerment. This will support PHNs in taking responsibility for their power and this will support raising their power and their ability to integrate health promotion into their practice.

Raising PHN self-efficacy

- The education, training, and practice of PHN skills are essential to maintaining the self-efficacy of PHNs.
- Empowering management practices need to involve sharing power, but to be effective must be perceived at an individual level to raise the self-efficacy of the PHN and this is something that can be achieved by practices like verbal feedback. This needs to include supporting making visible PHNs non-quantifiable work and fostering a seat at the table of power for PHNs.
- Provide greater and ongoing support and feedback to PHNs and ongoing efforts to raise the profile of the role.
- The leadership style provided to PHNs needs to be more democratic.
- Efforts at maintaining the self-efficacy of PHNs need to be maintained to foster the empowerment of PHNs.

The outcome is PHNs feels supported to act on their self-efficacy and ability and advocate for their role and will be better positioned to provide reciprocal support to their clients.

Recommendations for PHNs

PHN role and workload

- Engagement in action research to begin to define the PHN role with clarity.

When the role is clearer it will assist PHNs in becoming open to the need for change in their role and the need for sharing responsibility. Then, PHNs will be able to make a concrete case to their managers and the Health Service regarding their caseloads. When PHN responsibility becomes shared, PHNs begin to have time to plan and deliver health care with the health promotion approach. Working with the health promotion approach will foster the PHNs health promotion skills and the PHN will be reciprocally empowered in the process.

PHN access to empowering structures (including management structure)

- Involvement in brainstorming work to begin to identify the resources and supports that they need to support their empowerment and their empowering role.

By identifying the resources and support needs of PHNs they can work in consultation with the Health Service in a process to support change. When PHNs have the resources they need, this may aid in raising their sense of empowerment, which in turn may increase their efforts to engage in health promotion practice.

Continued professional development of PHNs

- Engagement in individual and group reflective practice sets particularly on issues such as power and role.
- Completion of structured personal development plans that feed back into the Health Service.
- Upskilling of communication and listening skills of PHNs.

These recommendations will support PHNs to understand more clearly how power influences them in their role, particularly with regard to their perceptions of autonomy and empowerment and how it influences both health choices that they and their clients may have. This will support PHNs in taking responsibility for their power. This will also help PHNs to become open to approaching and analysing the dynamic of power as it relates to themselves and their clients in a fuller broader sense to meet the needs of the evolving role. This will support PHNs to foster their advocacy role by focusing on the power gap and working to positively influence the determinants of health on health choices, rather than focusing on providing information and sending referrals.

Raising PHN self-efficacy

- Advocate for the development of mechanisms both formal and informal to raise the profile of the PHNs role and their self-efficacy in their roles.

The leadership style provided to PHNs needs to become more democratic, particularly as if PHNs receive ongoing support and feedback the shared governance agenda will be supported. This has the potential to raise the profile of the PHN role and PHNs self-efficacy and empowerment and support PHNs in their health promotion role and in advocating for their role.

Recommendations for the health service/ DoHC

PHN role and workload

- The Health Service needs to engage in consultative partnerships with PHNs and their managers to enable the process of the clarification of the PHN role.
- The DoHC need to be mandated to complete and launch the NAMIC report.
- The Health Service needs to provide the resources and funding PHNs need for their professional development to meet the needs of the health promotion approach.

Clear role and appropriate workloads will assist PHNs to remain open to the need for change in the role and the need for sharing responsibility and this will give PHNs time to plan and foster their skills in health promotion.

PHN access to empowering structures (including management structure)

- The Health Service needs to commit itself to advocacy for PHNs including addressing the PHNs need for leadership and support.
- The Health Service needs to introduce systems of critique of management structures and styles to ensure they meet the growth and development needs of PHNs and their clients.
- The HSE needs to develop management development strategy to include personal development planning for PHNs and their managers that feed back into the HSE.

- The HSE needs to commit to shared governance for PHNs.

Continued professional development of PHNs

- Facilitate the provision of properly sequenced training modules on power, empowerment, and health promotion.
- Encourage Managers to encourage PHNs to maintain personal development plan that feeds back into the HSE.
- Facilitate PHNs engagement in individual and group reflective practice sets to reflect on power to foster the empowerment role.
- Identify their training needs in relation health promotion.

The advocacy role needs the resource of time and appropriate education to support PHNs to reflect on power and make its effects clear and to support addressing the power gap. Intercommunity knowledge of PHNs is fundamental to their advocacy role. Relinquishing power is needed at all levels to support the client centred approach. This will support PHNs in taking responsibility for their power and this will raise their power.

Raising PHN self-efficacy

- Foster the development of mechanisms both formal and informal to raise the profile of the PHNs role and their self-efficacy in their roles.

Self-efficacy needs to be maintained on an ongoing basis by efforts directed at fostered the ability and self-efficacy of PHNs. PHNs need shared governance and true autonomy over their practice and freedom to practice autonomously. When the profile of the PHNs role and her self-efficacy are raised, she is supported to act on her ability. This leads to the PHN playing a leadership role in working to raise the self-efficacy of their clients.

Recommendations for educational institutes

Continued professional development of PHNs

- Facilitate and foster reflective practice time and reflective sets on a variety of issues such as professional competence, professional confidence, power, empowerment and the role of PHN.
- Work simultaneously with the HSE and in particular, PHN Managers in the development of properly sequenced CPD programmes to meet the needs of the evolving role.
- Prioritise health promotion within PHN pre-service and career development.
- Facilitate multi-disciplinary training.

9. References

Argyris, C. (1964) Integrating the individual and the organisation. London: Wiley

Armstrong, S. (2000) 'Public Health Nursing' in Robin, J (ed) (2000) *Nursing and Midwifery in Ireland in the twentieth Century.50 years of An Bord Altranais- Nursing Board 1950 – 2000.* Dublin: An Bord Altranais

Aston, M., Meagher-Stewart, D., Sheppard-Lemoine, D., Vukic, A. & Chirop, A (2006). Family Health Nursing and Empowering Relationships. *Paediatric Nursing* Jan/ Feb, 61- 67.

Barden, A.M, Griffin, RN, Quinn, M.T., Donahue, M, Fitzpatrick, J.J. TIDY (2011) Shared Governance and Empowerment in Registered Nurses working in a hospital setting, *Nursing Administration Quarterly*, Vol 35 (3) July/ Sep 2011, pp 212-218

Begley, C.M., Brady, A-M., Byrne, G., Macgregor, C., Griffiths.C. and Horan, P. (2004*) A study of the role and the workload of the Public Health Nurse in the Galway Community Care Area.* Dublin: University of Dublin, Trinity College

Benner, P. (1984) *From novice to expert. Excellence and Power in clinical nursing practice*. California: Addison Westley

Bowers, B (2011) Managing change by empowering staff *Nursing Times,* Vol 107 August 2011, p 19-21

Burnard, P. (1991) A method of analysing interview transcripts in qualitative research. *Nurse Education Today*, 11, 461-466

Bush, T. (1995) *Theories of Educational Management* 2nd Edition. University of Leicester. Paul Chapman Publishers

Business Innovative Insider (2006) 'Dealing with innovation overload' available: http://www.businessinnovationinsider.com/2006/08/dealing_with_innovation_overload/ (assessed 16th Sep 2006)

Byrd, M.E. (1995) A concept analysis of home visiting. *Public Health Nursing* 12(2) p 83-89

Carey, P. (2000). Community Health and Empowerment in Kerr, J., (ed), *Community Health Promotion: Challenges for Practice* p 27 – 44. Royal College of Nursing. London: Balliere Tindall:

Cawley, T & Mannix McNamara, P (2011) Public Health Nurse Perceptions of Empowerment and Advocacy in Child Health Surveillance in West Ireland Public Health Nursing. *Public Health Nursing* March/ April 28 (2) p 150-158

Chandler, G. (1991) Creating an environment to empower nurses. *Nursing Management* 22(8) p 20- 23

Chang, L (2008) Employee empowerment, innovative behaviour and job productivity of public health nurses: A cross-sectional questionnaire survey. *International Journal of Nursing Studies* Volume 45, Issue 10, October 2008, pp 1442-1448

Chavasse, J. (1992) New dimensions of employment in Nursing and challenges. *Journal of Advanced Nursing*, 17, p 1 - 2

Clarke, H.F. & Mass, H (1998) Comox Valley Nursing Centre: from collaboration to empowerment. *Public Health Nursing*. Volume. 15 No. 3. pp 216-224

Clarke, J. (2004) Public Health Nursing in Ireland: A Critical Overview. *Public Health Nursing* Vol. 21, No 2, pp 191 - 198

Cohen, B.E., and Reutter, L (2007). Development of the role of pubic health nurses in addressing child and family poverty: a framework for action. *Journal of Advanced Nursing,* 60 (1), 96-107.

Cohen, L., Manion, L. and Morrison, K. (eds) (2003) *Research methods in education*. 5th Edition. London: Routledge Falmer

Conger, J. and Kanungo, R. (1988). The Empowerment Process: Integrating Theory and Practice. *Acadamy of Management Review*, 1988, vol. 13. No 3, p 471 – 482

Department of Health (1997) *Public Health Nursing: a review*. Dublin: Department of Health

Department of Health (1998) *Commission on Nursing Report*. http://www.nursingboard.ie/Nursingcommission/Chp08.htm (last accessed 5th July 2005)

Department of Health and Children (2000) *Job description of the Public Health Nurse. Circular 41/2000*. Dublin: Department of Health and Children

Department of Health and Children (2001) *The Primary Care- A New Direction*. Dublin: Stationary House

Department of Health and Children (2003) *Nurses and Midwives understanding and experiences of empowerment in Ireland: final report*. Scott, P.A. and Matthews, A and Corbally, M. DCU: DoHC. September 2003

Edgecombe, G (2005) Healthy communities: is there a role for nurses/ nursing? In Contexts of Nursing (Daly, J., Speedy, S., and Jackson, D. eds) Churchill Livingstone Elsevier, Sydney p252-267

Elmore, R. F (1989) Backward mapping. in Moon, B., Murphy, P. and Raynor, J. (eds) *Policies for the Curriculum.* London: Open University Press

English National Board (ENB) (1991) *Framework for continuing professional education of Nurses, Midwives and Health visitors: guide to implementation.* London: ENB

Falk-Rafael, A. (2001) Empowerment as a process of evolving consciousness: a model of empowered caring. *Advances in Nursing Science*. 24(1) : p 1 – 16

Falk-Rafael, A. (2004) The effectiveness of Feminist Pedagogy in Empowering a Community of Leaders. *Journal of Nursing Education*, Vol 43, No. 3

Falk-Rafael, A. (2005) Speaking truth to power- Nursing's legacy and moral imperative.

Foucault, M. (1978) *The Archaeology of knowledge*. London: Tavistock

Foucault, M. (1980) *Power/ Knowledge: Selected Interviews and Other Writings 1972- 1977* C Gordan (ed). London: Harvester Wheatsheaf

Foucault, M. (1994). *The Cambridge companion to Foucault*. Gutting, Gary (ed) Cambridge: New York: Cambridge University Press.

Freire, P. (1972) *Cultural action for freedom*. London: Penguin Education

Fulton, Y. (1997). Nurses' views on empowerment: a critical social theory perspective. *Journal of Advanced Nursing,* 1997, 26, p 529- 536.

Gibson, C (1991) A concept analysis of empowerment. *Journal of Advanced Nursing*. 16.3. p 354-361

Gilbert, T. (1995) Nursing: empowerment and the problem of power. *Journal of Advanced Nursing*, 21, p 965- 871

Hackman, J.R. and Oldham, G.R. (1975) Development of the job diagnostic survey. Journal of Applied Psychology 60 (2) p 159-170

Hanafin, S. (1997b). The role of the Irish PHN: manager, clinician and health promoter. *Health visitor,* volume 8; p 295 - 297

Hanafin, S., Houston. and Cowley, S. (2002) Vertical equity in service provision: a model for the Irish public health nursing service. *Journal of Advanced Nursing*. 39 (1) p 68 - 76

Haugh, E, and Spence Laschinger, H. K. (1996) Power and opportunity in public health nursing work environments. *Public health nursing*. Vol. 13, no 1 p 42 - 49

Hawks, J., and Hromek, C. (1992) Nursing Practicum: Empowerment strategies. *Nursing Outlook.* 40, 5 p 231-234

Henderson, DJ (1995) Consciousness raising in participatory research. Method and methodology for emancipatory nursing inquiry. *Advances in Nursing Science* 17 p58-69

Illich, I. (1970) *Celebration of awareness: A call for institutional Revolution*

Jackson S., Baron, V., Cleverly, S., Dixon, J., Ferguson, J., Jackson, L., Hyewon Lee-Han, D.J., Onyalo, S., Uhera, P and Yoan, L.(1996) *Public health practitioners' perspectives on empowerment: definition, strategies and indicators.* New York, ON Toronto, Canada: North Youk Community Health promotion research unit.

Kanter, R.M. (1977) *Men and women of the Corporation*. 2n. New York: Basic Books

Kanter, R.M. (1983) *The change masters*. New York: Simon and Sonuster

Kanter, R.M. (1993) *Men and women of the Corporation*. 2nd Edition. New York: Basic Books

Kumar, S. (1999). Grounded theory. in Carter, Y., Shaw, S., Thomas, C. (eds*) An Introduction to Qualitative Methods for Health Professionals.* Royal College of General Practitioners.

Kuokkanen, L Leino- Kilpi, H. (2000).Power and empowerment in Nursing: Three theoretical approaches. *Journal of Advanced Nursing*; Jan 2000, Vol. 31 Issue 1, p 235

Kuokkanen, L., Leino- Kilpi, H. and Katajisto, J. (2003). Nurse empowerment, Job- Related Satisfaction and Organisational Commitment. *Journal of Nursing Care Quality*, Jul-Sep, Vol. 18. No. 3, p 184-192

Labonte, R (1994) *Health promotion and empowerment: Practice frameworks.* Toronto: Centre for Health Promotion. University of Toronto

Laschinger, H.K. and Wong, C. (1999) Staff Nurse empowerment and collective accountability: Effect on perceived productivity and self-rated work effectiveness. *Nursing Economics*, Nov/ Dec 1999, Vol 19, No.6.

Laschinger, H.K. S., Finnegan, J. and Shamian, J. (2001) The impact of workplace empowerment, organisational trust on staff nurses work satisfaction and organisational commitment. *Healthcare Management Review*.26(3), p 7 - 23

Laschinger, H.K.S., Wong, C., McMahon, L., and Kaufmann, C. (1999) Leader member behaviour on staff nurse empowerment, job tension and work effectiveness. *Journal of Nursing Administration* 29, p 28-39

Leahy- Warren, P. (1998) *Community Nursing- International perspective. A report for the Commission on Nursing*. Dr Geraldine McCarthy- Haslam (ed). Dublin: Stationary house

Lewin, M and Urmston, J. (2000). Flogging the dead horse: the myth of nursing empowerment? *Journal of Nursing Management* 8, p 209

Liou, C.Y. (2000) *The discrepancy between the competence of public health nurses and public health nursing practice*. Unpublished Masters Dissertation. University ofYang-Ming, Taipei, Taiwan

Malin, N. and Teaside, K. (1991) Caring versus empowerment considerations for nursing practice. *Journal of Advanced Nursing* 16 (3) p 657- 662

Maslow, A. H. (1968). *Towards a psychology of being.* 2nd Edition. London: Van Nostrand Reinfold Company

McDermott, K., Heather, K., Spence Laschinger, H.K. and Shamian, J. (1996). Work Empowerment and Organisational Commitment. *Nursing Management* Vol 27, No. 5, p 44 – 47.

McNeeseSmith, D.K. (2001) Staff nurse views of their productivity and nonproductivity. Health Care Management Review 26(2) p 7-19

Meagher-Stewart, D, Underwood, J, MacDonald, M, Schoenfeld, B, Blythe, J, Knibbs, K, Munroe, V, Lavoie-Tremblay, M, Ebrlich, A, Ganann, R Organizational Attributes that assure optimal utilization of Public Health Nurses. *Public Health Nursing* Vol.27 no.5 pp 433-441

Mullins, L.J.(1999) *Management and organisational behaviour.* 5th Edition. London: Prentice Hall

Nic Philibin, C, Griffiths, C., Byrne, G., Horan, P., Brady, A-M & Begley, C (2010) The role of the public health nurse in a changing society. *Journal of advanced nursing* 66(4) p 743-752

Naidoo, J. and Wills, J.(2000) *Health promotion Foundations for practice.* 2nd Edition.London: Balliere Tindall

Ning, S., Zhong, H., Libo, W & Qiujie, L (2009) The impact of nurse empowerment on job satisfaction. *Journal of advanced Nursing* 65(12) p 2642-2648

O Halloran, S. (2000) 'Rethinking nursing knowledge through action research'. In McNiff, J., McNamara, G. and Leonard, D. (2000) *Action research in Ireland.* p 113-128

O Sullivan, T. (1995) *A service without walls.* Dublin: Institute of Public Administration

Rodwell, C.M. (1996). An analysis of the concept of empowerment. *Journal of Advanced Nursing,* 1996, 23, p 305 – 313.

Rose, J & Glass, N (2008) The importance of emancipatory research to contemporary nursing practice *Contemporary Nurse* Volume 29, May 2008, p 8-22

Ryan, D., Mannix McNamara, P., and Deasy, C. (2006) *Health promotion in Ireland- Principles, Practice and Research.* Dublin: Gill and Macmillan

Ryles, S. (1999) A concept analysis of empowerment: its relationship to mental health nursing. *Journal of Advanced Nursing*; Mar 99, Vol. 29 Issue 3.

Sabiston, J.A., and Laschinher, H.K.S. (1995) Staff nurse work empowerment and perceived autonomy:testing Kanter's theory of structural power in organization. *Journal of Nursing Administration* 25(9), p 42-50

Salvage, J. (1992) The new nursing: empowering patients or empowering nurses? *In Policy issues in Nursing* Robinson, J., Gray,. And Elkan, R, (eds) Open university press, Milton Keyes, p 9 - 23

Schon, D(1983) *The reflective practitioner.* New York: Basic Books

Scott, S.G. and Bruce, R.A. (1994) Determinants of innovative behaviour: a path model of individual innovation in the workplace. *Acadamy Management Journal* 37(3) p 580-607

Scottish Executive (2006) Visible, Accessible and Integrated Care. Report of the Review of Nursing in the Community in Scotland. *Scottish Executive* Edinburgh, Edinburgh

Seedhouse, D. (1986) *Health the foundations for achievement.* Chichester. Redwood Books, Great Britain: John Wiley & Sons:

Skelton, J. (1997) Coding and Analysing Data. In: Carter, Y., and Thomas, C. *Research Methods in Primary Care.* Oxford, U.K.: Radcliffe Medical Press.

Spence Laschinger, H. K. (2001) Promoting Nurses health: effect of empowerment on job strain and work satisfaction. *Nursing Economics*; Mar/ April 2001, Vol. 19, Issue 2

Spreitzer, G. M. (1995) Psychological empowerment in the workplace: Dimensions, measurement, and validation. *Academy of Management Journal.* 38(5) p 1442-1462

Spreitzer,G.M. (1996 Socila structural characteristics of psychological empowerment. *Acadamy of Management Journal* 39(2) p483-504

Stevens, P.E. and Hall, J.M (1992) Applying critical theories to nursing communities. *Public health Nursing* 9(1) p 2-9

Swihart, D (2006) *Shared Governance: A Practical Approach to reshaping professional nursing practice*, Marble Head, MA, Inc

Tengland, P(2007) Empowerment: A goal or a means for health promotion *Medicine, Health Care and Philosophy* Vol 10 pp 197-207

Underwood, J.M., Meagher-Stewart, D.M., Deber, R.B., Baumann, A.O., MacDonald, M.B. , Akhar-Danesh, N., Schoenfeld, B.M., Ciliska, D.K., Blythe, J.M., Lavoie-Tremblay, M., Ehrlich, A.s., Knibbs, K.M., and Munroe, V.J. (2009) Building Community and Public health Nursing Capacity: A Synthesis of the National Community Health Nursing Study, *Canadian Journal of Public Health*, Sep/ Oct Vol 100 No. 5

Whitney, J. (1994) *The Trust Factor*. New York: McGraw-Hill, p 235

Williamson, D.L. and Drummond, J (2000) Enhancing low-income parents' capacities to promote their children's health: education is not enough. *Public Health Nursing* 17(2) p 121-131.

Wilson, B., and Laschinger, H.K.S. (1994) Staff nurse perceptions of job empowerment and organizational commitment: a test of Kanter's theory of structural power in organization. *Journal of Nursing Administration* 24(45) p 39-47

Wittmann-Price, RA (2004) Emancipation in decision making in women's health care. *Journal of Advanced Nursing* 47 p 437-445

World Health Organisation (1986) *Ottawa Charter for health promotion*. Geneva:WHO

Yeh, C.C. (2000) *The relationship between team member per4ceived empowerment and personal innovativeness*. Unpublished Masters dissertation. University of Ming Chuan, Taipei, Taiwan

Zerwekh, J. V. (1992) The practice of empowerment and coercion by expert public health nurses. *The British Library*. Vol 24, No 2, Summer 1992

18

Disseminating an Evidence-Based Disease Self-Management Program for Older Americans: Implications for Diversifying Participant Reach Through Delivery Site Adoption

Matthew Lee Smith[1,2,*], Basia Belza[3,4], Mary Altpeter[5],
SangNam Ahn[2,6], Justin B. Dickerson[2] and Marcia G. Ory[2]
[1]College of Public Health, University of Georgia, Athens, GA,
[2]School of Rural Public Health, Texas A&M Health Science Center, College Station, TX,
[3]Health Promotion Research Center,
School of Public Health, University of Washington, Seattle, WA,
[4]Biobehavioral Nursing and Health Systems,
School of Nursing, University of Washington, Seattle, WA,
[5]UNC Institute on Aging, University of North Carolina, Chapel Hill, NC,
[6]School of Public Health, University of Memphis, Memphis, TN,
USA

1. Introduction

As life expectancy among the American population increases, healthcare costs associated with aging will continue to represent a growing proportion of overall national health expenditures (Shortell et al., 2010). Much of this cost is associated with burgeoning chronic disease burden experienced by our aging society. Over the past decade, public health initiatives have focused additional effort on preventive behaviors and self-management skills to offset or address chronic disease and enhance the availability and accessibility of community- and home-based health-related resources and services (United States Senate, n.d.). Increasing the delivery and utilization of health-related programs in various settings enables older adults to receive vital education and support necessary to modify health behaviors while becoming more physically and socially active and embracing self-care practices. Such programs have shown benefit to prolong healthy life years and reduce the high costs associated with emergency care and institutionalization (Centers for Disease Control and Prevention, 2011a). As such, policymakers and practitioners alike continue to seek efficient ways to implement evidence-based community interventions and improve delivery networks to facilitate lower cost and more personalized solutions for aging Americans (National Prevention Council, 2011; U.S. Department of Health and Human Services, 2000). Despite recent successes in the United States, many questions remain

---

[*] Corresponding Author

concerning which types of community settings are best matched with certain demographic profiles of seniors to most effectively deliver health-related programs on a grand scale.

Organized by components of the RE-AIM framework (i.e., Reach and Adoption), this chapter uses data collected during a nationwide dissemination of the Chronic Disease Self-Management Program (CDSMP) through the aging services network to: 1) illustrate the geographic dispersion of the CDSMP in the United States between 2006 and 2009; 2) describe CDSMP delivery site types in terms of their neighborhood characteristics; 3) describe the personal and neighborhood-level characteristics of older adults who enrolled in CDSMP by delivery site type; and 4) discuss policy and practice implications for disseminating community-based interventions to serve diverse populations of older Americans.

1.1 Health of an aging population

While many older adults report being healthy and independent, older Americans often report high rates of chronic illnesses that threaten that health and independence (U. S. Census Bureau, 2008). Going beyond the often–cited statistic that over 80 percent of older adults (65+) have at least one chronic illness (Centers for Disease Control and Prevention, 2011b; Martin, Freedman et al., 2010), there is growing attention to multiple co-morbidities and calls for a new multiple conditions framework for better understanding the health and societal burdens of chronic illnesses (Alliance for Health Reform, 2011; Interagency Workgroup on Multiple Chronic Conditions, 2010). Chronic conditions are a leading cause of death and account for more than three-quarters of all health expenditures in the United States (Goodman et al., 2004). A multi-pronged, multi-level approach is needed to address the prevention and management of chronic conditions and enable older adults to age more successfully (Centers for Disease Control and Prevention, 2011b). Effective strategies for: 1) reducing the onset or exacerbation of chronic conditions include building healthy communities to make it easy for residents to select the 'right' lifestyle choices (Ory et al., 2009; Satariano et al., 2011); 2) promoting stronger linkages between clinical and community resources (Bolin et al., 2011); 3) helping older adults develop skills to better communicate with healthcare providers (Ory, et al., 2006); and 4) empowering older adults to take a more active role in their healthcare through adoption of evidence-based self-management strategies (Hughes et al., 2011). Despite the well-documented importance of self-management for minimizing the burdens of chronic illnesses and disabilities (Centers for Disease Control and Prevention, 2011a; Ory, et al., 1998), many barriers to self-management have been identified among older adults including diminished mental and physical functioning, lack of knowledge or confidence in engaging in self-management behaviors, and minimal familial or community supports for engaging in such activities (Center for Healthy Aging, n.d.). Evidence-based interventions that address these types of barriers are essential to promoting healthy aging.

1.2 Chronic Disease Self-Management Program (CDSMP)

Over twenty years of research at Stanford University has resulted in a widely disseminated self-management program for people with chronic conditions, the Chronic Disease Self-Management Program (CDSMP). Developed by Dr. Kate Lorig and her colleagues at the Patient Education Research Center at Stanford University (http://patienteducation.stanford.edu), there have now been several rigorous clinical trials testing and replicating the original research (Lorig et al., 2002; Lorig et al., 2001; Lorig, Sobel, et al., 2001; Lorig et al., 1999).

As a result of attending CDSMP, health status and behavioral improvements have been reported for participants of diverse ages, cultures, and ethnicities. For examples, in demonstrating the evidence-based, positive outcomes were identified in terms of self-rated health, reduced disability, social and role activities, and health distress; reduced fatigue and pain symptomatology; increased physical activity; greater skill in coping strategies and symptom management; better provider-patient communications; and reduced healthcare utilization (Lorig, et al., 2002; Lorig, Ritter, et al., 2001; Lorig, Sobel, et al., 2001; Lorig, et al., 1999).

Based on social-cognitive theory, CDSMP has translated fundamental behavior science change principles into practice. As the program has developed, it has become well scripted so original findings can continually be reproduced in different settings and populations. Whether taught in New York or Texas, the workshops' structure will be similar because trained facilitators follow a detailed implementation manual (http://patienteducation.stanford.edu).

Classes are held for persons with chronic conditions in a small group setting (e.g., 10-16 participants) over a six week period for 2.5 hours each session. The workshop welcomes participants with all types of chronic diseases recognizing individuals suffering from any specific chronic condition face common problems. These common problems include: pain management, diet, sleep and fatigue, medications, exercise and communicating with clinicians. The workshop is designed to help participants develop skills at managing symptoms and learn coping strategies using well-tested behavior change strategies such as action planning and feedback, behavior modeling, problem-solving techniques, and decision making (Ory, et al., 2002). In recognition of the importance of outreach to potential participants who can benefit from increased self-management skills, the workshops are offered in multiple community settings such as senior centers, churches, libraries, and healthcare settings.

To maintain quality control, Stanford University requires certification and licensure for all parties who deliver CDSMP. The program developers offer a training program for Master Trainers to learn and develop skills necessary to train and supervise the Lay Leaders (class facilitators) and ensure the quality of local programs. Implementation protocols call for two trained Lay Leaders (i.e., at least one of the leaders is a non-health professional with a chronic disease) to co-facilitate the program. Ideally, Lay Leaders are matched to the CDSMP program participants by race/ethnicity, culture, gender, and/or age. Stanford has trained Master Trainers from all 50 states as well as countries around the world (e.g., Canada, Australia, China, Japan, Norway and other Scandinavian countries, and the United Kingdom).

Several aspects make CDSMP a unique evidence-based self-management program. Its benefits have been documented in multiple settings and populations. Its well-specified implementation manual assists in scalability and makes further dissemination easier and more structured. Additionally, the well-honed training infrastructure and continual feedback via fidelity checklists helps ensure program uniformity as it is disseminated in different locations.

1.3 The Administration on aging's evidence-based disease prevention initiative

In 2003, the Administration on Aging (AoA) began funding pilot programs to test the translation of the Evidence-Based Disease and Disability Prevention programs (EBDDP), including the CDSMP program, in the Aging Services Network's community-based settings. Based on the positive results of these pilot programs, the AoA increased its Federal support of EBDDP in 2006. A total of $4,542,300 USD was allocated between 16 states to support collaborations between the aging and public health networks at the state and local level to implement these programs for older adults. In 2007, $5,841,680 USD was allocated between eight additional states. The AoA continued supporting these grants through competitive supplements of $5,091,680 USD in 2008 and $5,091,680 USD in 2009. Since 2006, the AoA has awarded $22 million and leveraged an additional $20 million to support evidence-based programs in 27 States. This funding supported the development of a delivery infrastructure for evidence-based programs to serve older adults in various community-based settings. The AoA leads the EBDDP initiative in partnership with the Centers for Disease Control & Prevention (Centers for Disease Control and Prevention (CDC)), Agency for Healthcare Research & Quality (AHRQ), Centers for Medicare and Medicaid Services (CMMS), Health Resources & Services Administration (HRSA), Substance Abuse & Mental Health Services Administration (SAMHSA), and over 30 private foundations.

1.4 The role of the RE-AIM framework

To be effective and meet predetermined expectations, the CDSMP national roll-out necessitated a broad public health perspective. In consultation with translational researchers, AoA administrators drew upon the RE-AIM framework as an organizing framework for program planning and evaluation (Glasgow, et al., 2001; Glasgow, et al., 1999). As part of funding requirements, AoA grantees were expected to describe their use of the RE-AIM model to plan, implement, evaluate, and sustain their proposed health promotion programming.

Fig. 1. The RE-AIM Framework

The RE-AIM framework has been employed to encourage program planners, evaluators, researchers, funders, and policy-makers to heed the essential program elements that can improve the implementation, adoption, and sustainability of effective, evidence-based health promotion programs.

RE-AIM is a mnemonic that helps community practitioners focus on program *reach, effectiveness, adoption, implementation,* and *maintenance.*

In the following section, we briefly describe the different RE-AIM components (Administration on Aging, 2010; Glasgow, et al., 2001; Glasgow, et al., 1999).

Reach focuses on the extent to which a program reaches the intended target population. Monitoring *reach* is important to determine if the target audience is participating in the program, in what numbers, and the percentage of program completion and attrition. It also helps to determine the adequacy of marketing efforts, recruitment and retention of participants and whether certain program sites are having problems with filling workshops or attendance (retention).

Effectiveness focuses on whether a program is achieving the same participant outcomes and having the same impact as in the original research design. Monitoring *effectiveness* helps to assess whether a program is producing positive changes in participants' health and well-being and whether there are any unintended (positive or negative) outcomes. Demonstrating such program value and return on investment is important to key stakeholders.

Adoption focuses on the extent to which host agencies and implementation sites deliver and embed the program into routine activities and the level of organizational support that is provided. Monitoring *adoption* helps to determine whether there are enough partners, implementation sites, frequency of workshops and personnel to deliver the program and to reach the target population; how well partners and sites are supporting the programs; whether the settings are appropriate and accessible for those to be reached; and whether the program can “go to scale.”

Implementation focuses on fidelity monitoring, that is, the consistency of a program’s delivery in different settings and with different instructors. Fidelity monitoring may include assessing consistency with the intended program design, training, delivery, and participant mastery and application. Monitoring *implementation* helps to ensure the program is delivered in a quality manner, no matter how often, by whom or in what setting. This, in turn, helps to document that the participant outcomes can really be attributed to the program and can help to identify areas of need for improvement or changes in training or program delivery.

Maintenance focuses on assessing: 1) at the program level, the extent to which a program becomes institutionalized or part of the routine organizational practice and policy and, 2) at the individual level, the extent to which participants sustain long-term benefits from completing the program. Monitoring *maintenance* enables program managers to assess marketing effectiveness; expansion of accessibility to new partners, program sites and new populations and the capacity to scale state-wide; and the extent to which the program is sustained and embedded within the state’s evidence-based prevention program distribution and delivery system and other health and long-term supportive services systems.

1.5 An emphasis on reach and adoption

While all elements of the RE-AIM Framework are essential to develop, organize, deliver, and evaluate programs disseminated on a grand scale, this chapter specifically emphasizes the aspects of Reach and Adoption in this national AoA initiative. The authors elected to highlight these RE-AIM elements because they are indicative to the success of providing access to diverse populations. Below, we briefly describe our rationale in focusing on these two RE-AIM elements as a way of examining the relationship between program reach and delivery site. We see the types of participants reached as influenced by the types of agencies and organizations who adopt and deliver the program. Theoretically, there is an association between Reach and Adoption with a larger number of participants enrolling in a program if more organizations adopt the program and deliver it in close proximity to the participants' place of residence. Moreover, organizations and delivery site types typically service older adults with varying characteristics (e.g., age, race/ethnicity, residential rurality). A diversification of the types of delivery sites that offer the program may be needed to increase the likelihood that a diverse group of older adults are attracted to and enrolled in an evidence-based program. Additionally, some organizations and agencies are generally located in geographic areas with varying neighborhood-level characteristics. Thus, they attract different clientele, which is often influenced by the types of individuals who reside in the area.

Expanding the number of delivery sites (regardless of types) over a wide geographic community landscape may increase the chances that participants will have access to the program, enroll in the program, and complete the program. Offering the program in close proximity to the participants' residence reduces the distance they must drive (or have someone else drive them) to access the program or service. This follows the basic community health development principle that programs and services will be utilized more frequently if offered in areas where people live, work, play, and/or pray.

This chapter reports participant (i.e., Reach) and delivery site (i.e., Adoption) characteristics associated with data collected nationwide by Senior Services between September 2006 and August 2009 for the CDSMP program. Within each grantee state, a designated data manager collected participant-level demographics, participant-level program attendance, and course or workshop data (i.e., including the delivery site type and host organization). The data manager completed a standardized template provided by AoA (accompanied by a data dictionary and user guide) for completed workshops each 6-month reporting period. These de-identified data were sent to AoA, exported from the state's internal data management system, and compiled into a centralized master database. Senior Services personnel then performed data quality checks to validate the integrity of the data. Data errors were referred to the state data managers for correction.

Although CDSMP was delivered beyond this AoA-funded initiative, we only included data collected as part of these grant efforts. Additionally, in the 2006-2009 initiative AoA did not require the systematic collection of outcome data, given that CDSMP was a proven evidence-based program. Thus, health-related outcomes are not reported in this chapter.

2. Participants and procedures

The data collected during this initiative was geocoded with ArcGIS ArcMap 10 using the ArcGIS Online U.S. Streets Geocode Service. Of the 1,339 CDSMP delivery site addresses, 1,306 (97.5%) were geocoded to their street location. Participant records were geocoded using their residential ZIP Code. Of the 23,091 CDSMP participants, 16,356 (70.8%) were geocoded. Of the 6,735 participant cases that could not be geocoded, 75 (0.3%) had invalid ZIP Codes, and 6,660 (28.9%) had missing ZIP Code data.

The neighborhood-level demographic data (i.e., median income, percent of the population age 65 years and older, percent of the population that was non-Hispanic white) was obtained from the U.S. Census Bureau, Census 2000, Summary File. This information was linked to participant records based on the ZIP Code of the participants' residence and program delivery site address. Data used to determine rurality (i.e., metro versus non-metro), was based on the county in which the participant resided and/or delivery site was located and are from the United States Department of Agriculture's 2003 Rural-Urban Continuum Codes (United States Department of Agriculture, 2004).

2.1 Measures

To examine personal and neighborhood characteristics associated with CDSMP delivery sites during this nationwide dissemination effort, a variety of measures were used.

Delivery Site Types. Data pertaining to CDSMP delivery site types were gathered administratively, as previously described. Participant cases attending the five most prevalent delivery site types were compared in these analyses: senior centers or Area Agencies on Aging (AAA), residential facilities, healthcare organizations, community or multi-purpose centers, and faith-based organizations. Participants attending CDSMP at any other delivery site types were omitted because of inadequate case sizes. The following delivery site types were omitted from analyses: State and County Public Health Departments (n = 29, 0.3%); municipal government (n = 58, 0.5%), workplace (n = 95, 0.8%), and Parks Department facilities (n = 108, 0.09%). Further, delivery sites categorized as "other" (n = 802, 7.0%) were omitted because of the potential difficulty to interpret findings associated with this delivery site type.

Note: Area Agencies on Aging (AAA) were established under the Older Americans Act (OAA) to serve the needs of older Americans (i.e., aged 60 and older). This nationwide network of agencies spans all 50 of the United States in an effort to keep older adults residing independently in their homes while providing valuable social services and community-based programs (Bookman, et al., 2007).

Neighborhood Characteristics. Using participants' residential ZIP Codes, geographic information system (GIS) software was used to generate neighborhood-level variables for each participant. Neighborhood characteristics included residential rurality (i.e., metro residence [urban] or non-metro [rural] residence based on the rural-urban commuting area codes (RUCA)), median household income for residents residing in the participants' ZIP code (i.e., interpreted in increments of $10,000 USD), the percent of residents aged 65 years and older residing in the participants' ZIP code, and the percent of non-Hispanic white residents residing in the participants' ZIP code. Using organizational ZIP codes, geographic information system software was used to generate neighborhood-level variables for each delivery site (i.e., site rurality, median household income, percent of residents aged 65 years and older, and percent of non-Hispanic white participants).

Personal Characteristics. Personal characteristics of the participants included age (i.e., ranging from 51 to 103 years), sex (i.e., male, female), race/ethnicity (i.e., non-Hispanic white, African

American or Black, Hispanic or Latino, Other/Multiple races), and the highest level of education received (i.e., less than high school, graduated high school, more than high school). Participants also self-reported their living situation (i.e., lives alone, lives with others).

2.2 Data analysis

For the purposes of this study, analyses were limited to participants who enrolled in CDSMP. Additionally, participant cases were omitted for those who attended programs hosted at delivery sites other than the five most prevalent sites noted above. All statistical analyses for this study were performed using SPSS (version 17). Frequencies were calculated for all major study variables which were examined in relationship to the program delivery site type. Frequency distribution differences for categorical variables were assessed using Pearson's chi-squared tests. One-way analyses of variance (*f* statistics) were used to identify mean differences between program delivery site types for continuous variables. Multinomial logistic regression was used to identify personal characteristics and participants' neighborhood-level characteristics associated with the type of delivery site they attended (i.e., senior center or Area Agency on Aging sites served as the referent group).

3. Nationwide dissemination of CDSMP

Our initial goal is to depict the geographic dispersion of CDSMP delivered nationwide via the AoA initiative between 2006 and 2009 in the five most prevalent types of delivery sites. As indicated in Figure 1, over this funding period, a total of 1,339 CDSMP workshops were delivered across 26 states. Nationwide, 358 (26.7%) workshops were delivered at senior centers/AAA, 271 (20.2%) at residential facilities, 195 (14.6%) at healthcare organizations, 142 (10.6%) at community or multi-purpose centers, and 131 (9.8%) at faith-based organizations.

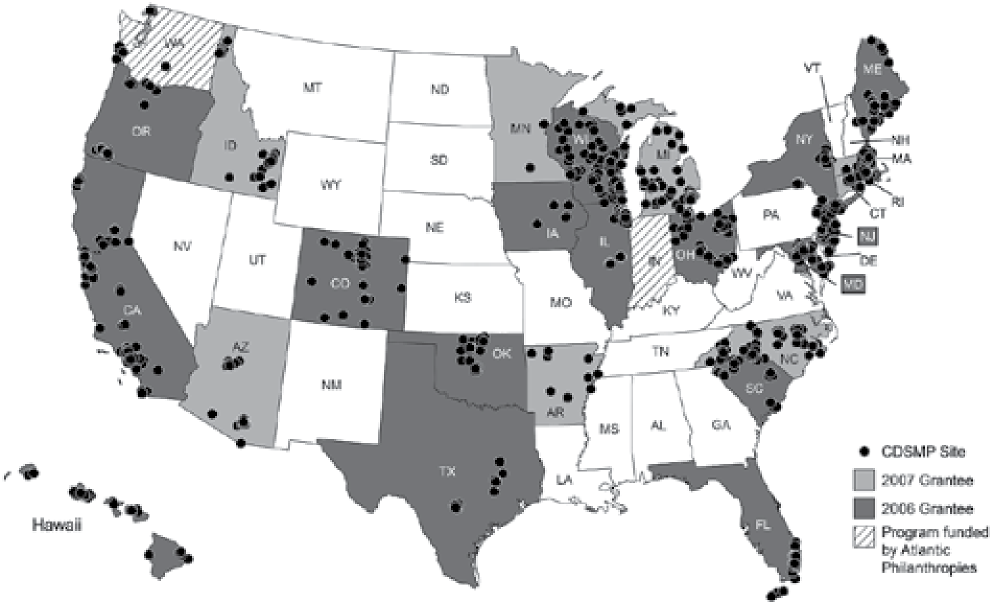


Fig. 2. National CDSMP Delivery Site Dispersion: 2006 to 2009

Patterns of CDSMP workshop delivery differed within states and could be classified as concentrated regionally (as in Texas and Florida) or dispersed statewide (as in Wisconsin and Michigan). The delivery dispersion of CDSMP was influenced by the AAA infrastructure, funding timeline, and allocation within each state. Generally speaking, those states funded in earlier years of the initiative delivered a larger number of CDSMP workshops.

3.1 Neighborhood characteristics of CDSMP delivery site types

A major research and practice question is the relationship between CDSMP delivery sites and socio-demographic neighborhood characteristics. As indicated in Table 1, 84% percent of participants attended CDSMP workshops delivered in metro (or urban) areas. On average, participants attended workshops delivered in ZIP codes where the median household income was $56,700 USD (±$31,700) and in areas where 13% (±6%) of the population was age 65 years and older. On average, CDSMP participants attended workshops delivered in ZIP codes comprised of 71% (±26%) non-Hispanic white residents.

When comparing these neighborhood characteristics by delivery site type, a larger proportion of CDSMP workshops in non-metro (or rural) areas were delivered in senior centers/AAA (19.9%) and faith-based organizations (18.2%) compared to community/multi-purpose centers (13.8%) and residential facilities (11.0%). On average, workshops at healthcare organizations were delivered in more affluent areas (median household income of $71,400) whereas workshops at residential facilities were delivered in less affluent areas (median household income of $47,200). On average, CDSMP workshops at community/multi-purpose facilities (64% non-Hispanic white) were delivered in more racially/ethnically diverse areas compared to workshops offered at residential facilities (75% non-Hispanic white) and senior centers/AAA (73% non-Hispanic white).

3.2 Personal characteristics of CDSMP participants by delivery site type

Another primary goal was to understand if there was a relationship between delivery site type and participant characteristics. Personal characteristics of study participants are also presented in Table 1. Of the 10,242 CDSMP participants with complete data on all variables of interest, 34.8% (n = 3,562) attended workshops at senior centers/AAA, 18.2% (n = 1,860) at residential facilities, 26.7% (n =2,738) at healthcare organizations, 10.8% (n = 1,108) at community or multi-purpose centers, and 9.5% (n = 974) at faith-based organizations.

Overall, the average age of participants was 72 years (±9.71). The majority of CDSMP participants were female (78.8%), non-Hispanic white (62.8%), lived with others (53.9%), and resided in metro areas (84.3%). Over 44% of participants had more than a high school education, whereas 28.0% graduate high school and 27.5% had less than a high school education. On average, participants resided in ZIP codes where the median household income was $43,000 USD (±$15,400) and in areas where 13% (±0.6%) of the population was age 65 years and older. On average, CDSMP participants resided in ZIP codes comprised of 70% (±26%) non-Hispanic white residents.

Delivery Site Characteristics	SC/AAA (n = 6441)	RF (n = 4523)	HO (n = 5809)	C/MPC (n = 3072)	FBO (n = 2168)	Total (n = 22013)	X^2 or t	P
Metro (Delivery Site)	5161 (80.1%)	4024 (89.0%)	4889 (84.2%)	2648 (86.2%)	1773 (81.8%)	18495 (84.0%)	174.18	<0.001
Non-Metro (Delivery Site)	1280 (19.9%)	499 (11.0%)	920 (15.8%)	424 (13.8%)	395 (18.2%)	3518 (16.0%)		
Median Income for ZIP: Delivery Site	5.02 (±2.98)	4.72 (±2.94)	7.14 (±2.99)	5.95 (±3.31)	5.54 (±3.02)	5.67(± 0.17)	537.12	<0.001
Percent of Delivery Site ZIP Population: Age 65+	0.14 (±0.05)	0.15 (±0.06)	0.12 (±0.05)	0.13 (±0.06)	0.13 (±0.05)	0.13 (±0.06)	199.29	<0.001
Percent of Delivery Site ZIP Population: Non-Hispanic White	0.73 (±0.26)	0.75 (±0.24)	0.68 (±0.25)	0.64 (±0.27)	0.71 (±0.25)	0.71 (±0.26)	127.69	<0.001

Participant Characteristics	SC/AAA (n = 3562)	RF (n = 1860)	HO (n = 2738)	C/MPC (n = 1108)	FBO (n = 974)	Total (n = 10242)	X^2 or t	P
Age	73.35 (±8.90)	76.39 (±9.86)	69.03 (±9.29)	70.54 (±9.97)	70.52 (±9.45)	72.17 (±9.71)	201.83	<0.001
Male	695 (19.5%)	251 (13.5%)	809 (29.5%)	239 (21.6%)	177 (18.2%)	2171 (21.2%)	191.83	<0.001
Female	2867 (80.5%)	1609 (86.5%)	1929 (70.5%)	869 (78.4%)	797 (81.8%)	8071 (78.8%)		
Less than High School	709 (19.9%)	458 (24.6%)	1023 (37.4%)	392 (35.4%)	235 (24.1%)	2817 (27.5%)	306.08	<0.001
Graduated High School	1095 (30.7%)	612 (32.9%)	636 (23.2%)	261 (23.6%)	262 (26.9%)	2866 (28.0%)		
More than High School	1758 (49.4%)	790 (42.5%)	1079 (39.4%)	455 (41.1%)	477 (49.0%)	4559 (44.5%)		
Non-Hispanic White	2274 (63.8%)	1253 (67.4%)	1810 (66.1%)	554 (50.0%)	536 (55.0%)	6427 (62.8%)	309.13	<0.001
Hispanic	356 (10.0%)	273 (14.7%)	419 (15.3%)	195 (17.6%)	130 (13.3%)	1373 (13.4%)		
African American	477 (13.4%)	182 (9.8%)	322 (11.8%)	125 (11.3%)	148 (15.2%)	1254 (12.2%)		
Other / Multiple Races	455 (12.8%)	152 (8.2%)	187 (6.8%)	234 (21.1%)	160 (16.4%)	1188 (11.6%)		
Lives with Others	1804 (50.6%)	462 (24.8%)	1950 (71.2%)	683 (61.6%)	622 (63.9%)	5521 (53.9%)	1043.56	<0.001
Lives Alone	1758 (49.4%)	1398 (75.2%)	788 (28.8%)	425 (38.4%)	352 (36.1%)	4721 (46.1%)		
Metro (Participant)	2805 (78.7%)	1680 (90.3%)	2407 (87.9%)	959 (86.6%)	778 (79.9%)	8629 (84.3%)	179.08	<0.001
Non-Metro (Participant)	757 (21.3%)	180 (9.7%)	331 (12.1%)	149 (13.4%)	196 (20.1%)	1613 (15.7%)		
Median Income for ZIP: Particpant Residence	4.34 (±1.45)	4.18 (±1.59)	4.64 (±1.58)	4.27 (±1.65)	4.05 (±1.39)	4.30 (±1.54)	40.18	<0.001
Percent of ZIP Population: Age 65+	0.13 (±0.05)	0.14 (±0.06)	0.12 (±0.06)	0.13 (±0.06)	0.13 (±0.05)	0.13 (±.006)	56.96	<0.001
Percent of ZIP Population: Non-Hispanic White	0.71 (±0.27)	0.74 (±0.24)	0.69 (±0.25)	0.66 (±0.26)	0.70 (±0.25)	0.70 (±0.26)	21.29	<0.001

*SC/AAA = Senior Center/Area Agency on Aging; RF = Residential Facility; HO = Healthcare Organization; C/MPC = Community/Multi-purpose Center; FBO = Faith-Based

**Means, standard deviations, and t-tests reported for continuous variables

Table 1. Delivery Site and Participant Characteristics by Delivery Site Type

To fully gauge how participant Reach differs by delivery site type relative to those enrolled at senior centers/AAA, we must first briefly describe the profile of participants who attended CDSMP workshops at senior centers/AAA. On average, these participants were age 73 years (±8.90). The majority of these CDSMP participants were female (80.5%), non-Hispanic white (63.8%), lived with others (50.6%), and resided in metro areas (78.7%). Over 49% of CDSMP participants who attended workshops at senior centers/AAA had more than a high school education, whereas 30.7% graduate high school and 19.9% had less than a high school education. On average, participants resided in ZIP codes where the median household income was $43,400 USD (±$14,500) and in areas where 13% (±5%) of the population was age 65 years and older. On average, CDSMP participants resided in ZIP codes comprised of 71% (±27%) non-Hispanic white residents.

When comparing participant characteristics by delivery site type, significant variation was observed relative to those who attended CDSMP workshops at senior centers/AAA. Residential facilities enrolled an older group of participants (76 years ±9.86), whereas healthcare organizations enrolled a younger group of participants (69 years ±9.29). The largest proportion of male workshop participants was reached in healthcare organizations (29.5%), whereas the smallest proportion was reached in residential facilities (13.5%). A larger proportion of participants with less than a high school education were reached in healthcare organizations (37.4%) and community or multi-purpose centers (35.4%). The greatest racial and ethnic diversity among CDSMP participants was observed among workshops delivered at community or multi-purpose centers and faith-based organizations. The largest proportion of participants living alone attended workshops at residential facilities (75.2%), whereas the smallest proportion of attended workshops at healthcare organizations (28.8%). While the largest proportion of rural-residing residents were reached in senior centers/AAA (21.3%), faith-based organizations also reached a larger proportion of these participants (20.1%), especially compared to the proportion of rural-residing residents reached by residential facilities (9.7%). CDSMP participants who attended workshops at faith-based organizations resided in the least affluent areas ($40,500 USD ±$13,900), whereas those who attended workshops at healthcare organizations resided in the most affluent areas ($46,400 USD ±$15,800). Participants who attended workshops at community or multi-purpose centers resided in the most racially/ethnically diverse areas (66% non-Hispanic white residents ±26%).

3.3 Delivery site type profiles by neighborhood characteristics

Little is known about the relationship between neighborhood-level characteristics of where CDSMP workshops were attended and delivery site type (i.e., Adoption characteristics). In this section, we will briefly describe each neighborhood-level delivery site type profile relative to workshops attended at senior centers/AAA. Using multinomial logistic regression, we fit a model with delivery site type as the dependent variable (attending CDSMP workshops at a senior center/AAA served as the referent group) and delivery site neighborhood characteristics were entered as independent variables. Table 2 contains the results of this analysis.

Residential Facilities. Compared to workshops attended at senior centers/AAA, participants were less likely to attend workshops delivered at residential facilities in rural areas (OR = 0.440, P < 0.001). Participants who attended workshops at residential facilities did so in areas that were less affluent (OR = 0.981, P = 0.002) and had larger proportions of the population who were non-Hispanic white (OR = 1.648, P < 0.001) and age 65 and older (OR = 8.942, P < 0.001).

	RF				HO				C/MPC				FBO			
			95% CI				95% CI				95% CI				95% CI	
	P	O.R.	Lower	Upper	P	O.R.	Lower	Upper	P	O.R.	Lower	Upper	P	O.R.	Lower	Upper
Non-Metro (Delivery Site)	<0.001	0.440	0.392	0.494	0.001	0.839	0.758	0.930	<0.001	0.786	0.693	0.892	0.875	0.989	0.867	1.130
Metro (Delivery Site)	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Median Income for Delivery Site ZIP	0.002	0.981	0.969	0.993	<0.001	1.256	1.240	1.272	<0.001	1.090	1.075	1.105	<0.001	1.048	1.031	1.064
Percent of Delivery Site ZIP Population: Age 65+	<0.001	8.942	4.478	17.859	<0.001	0.002	0.001	0.005	<0.001	0.167	0.069	0.402	<0.001	0.009	0.003	0.026
Percent of Delivery Site ZIP Population: Non-Hispanic White	<0.001	1.648	1.391	1.953	0.738	0.974	0.832	1.139	<0.001	0.362	0.305	0.430	0.225	1.139	0.923	1.405

*SC/AAA = Senior Center/Area Agency on Aging; RF = Residential Facility; HO = Healthcare Organization; C/MPC = Community/Multi-purpose Center; FBO = Faith-Based Organization

**Referent Group: Senior Centers / Area Agency on Aging

n = 22,013 (Nagelkerke = 0.138)

Table 2. Neighborhood Characteristics of the Delivery Site by Delivery Site Type

Healthcare Organizations. Compared to workshops attended at senior centers/AAA, participants were less likely to attend workshops delivered at healthcare organizations in rural areas (OR = 0.839, P = 0.001). Participants who attended workshops at healthcare organizations did so in areas that were more affluent (OR = 1.256, P < 0.001) and had smaller proportions of the population who were age 65 and older (OR = 0.002, P = 0.001).

Community or Multi-purpose Centers. Compared to workshops attended at senior centers/AAA, participants were less likely to attend workshops delivered at community or multi-purpose centers in rural areas (OR = 0.786, P < 0.001). Participants who attended workshops at community or multi-purpose centers did so in areas that were more affluent (OR = 1.090, P < 0.001) and had smaller proportions of the population who were non-Hispanic white (OR = 0.362, P < 0.001) and age 65 and older (OR = 0.167, P < 0.001).

Faith-Based Organizations. Compared to workshops attended at senior centers/AAA, participants who attended workshops at faith-based organizations did so in areas that were more affluent (OR = 1.048, P < 0.001) and had smaller proportions of the population who were age 65 and older (OR = 0.009, P < 0.001).

3.4 Participant profiles by personal and neighborhood characteristics

We also have limited information on the relationship between participant characteristics and type of CDSMP delivery site attended (i.e., participant Reach). In this section we will briefly describe each participant-level delivery site type profile relative to workshops attended at senior centers/AAA. Using multinomial logistic regression, we fit a model with delivery site type as the dependent variable (attending CDSMP workshops at a senior center/AAA served as the referent group) and participants' personal and neighborhood characteristics were entered as independent variables. Table 3 contains the results of this analysis.

Residential Facilities. Compared to workshops attended at senior centers/AAA, participants who attended workshops delivered at residential facilities were more likely to be older (OR = 1.023, P < 0.001) and female (OR = 1.242, P = 0.010). These individuals were less likely to have more than a high school education (OR = 0.771, P = 0.001) and less likely to be African American (OR = 0.664, P < 0.001) or of other/multiple races (OR = 0.760, P = 0.017). Participants who attended workshops delivered at residential facilities were more likely to live alone (OR = 2.644, P < 0.001) and less likely to reside in rural areas (OR = 0.362, P < 0.001). These participants also resided in areas that were less affluent (OR = 0.877, P < 0.001) and had larger proportions of the population who were non-Hispanic white (OR = 1.568, P = 0.006) and age 65 and older (OR = 4.097, P = 0.008).

Healthcare Organizations. Compared to workshops attended at senior centers/AAA, participants who attended workshops delivered at healthcare organizations were less likely to be older (OR = 0.950, P < 0.001) and female (OR = 0.634, P = 0.010). These individuals were less likely to have graduated from high school (OR = 0.424, P < 0.001) or have more than a high school education (OR = 0.372, P < 0.001). They were less likely to be Hispanic (OR = 0.701, P < 0.001), African American (OR = 0.578, P < 0.001), or of other/multiple races (OR = 0.309, P < 0.001). Participants who attended workshops delivered at healthcare organizations were less likely to live alone (OR = 0.518, P < 0.001) and reside in rural areas (OR = 0.648, P < 0.001). These participants also resided in areas that were more affluent (OR = 1.138, P < 0.001) and had smaller proportions of the population who were non-Hispanic white (OR = 0.567, P < 0.001) and age 65 and older (OR = 0.118, P < 0.001).

	RF				HO				C/MPC				FBO			
			95% CI				95% CI				95% CI				95% CI	
	P	O.R.	Lower	Upper	P	O.R.	Lower	Upper	P	O.R.	Lower	Upper	P	O.R.	Lower	Upper
Age	<0.001	1.023	1.016	1.030	<0.001	0.950	0.945	0.956	<0.001	0.970	0.963	0.978	<0.001	0.977	0.969	0.985
Female	0.010	1.242	1.054	1.464	<0.001	0.634	0.559	0.718	0.423	0.932	0.786	1.106	0.083	1.181	0.979	1.424
Male	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
More than High School	0.001	0.771	0.658	0.903	<0.001	0.372	0.324	0.428	<0.001	0.526	0.441	0.628	0.047	0.820	0.674	0.997
Graduated High School	0.141	0.884	0.750	1.042	<0.001	0.424	0.365	0.492	<0.001	0.495	0.409	0.600	0.008	0.754	0.612	0.930
Less than High School	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Other / Multiple Races	0.017	0.760	0.607	0.952	<0.001	0.309	0.250	0.383	<0.001	1.666	1.338	2.074	<0.001	1.725	1.361	2.188
African American	<0.001	0.664	0.533	0.828	<0.001	0.578	0.477	0.701	0.097	0.806	0.625	1.040	0.016	1.360	1.060	1.744
Hispanic	0.075	1.194	0.982	1.453	<0.001	0.701	0.583	0.843	0.008	1.356	1.084	1.697	0.230	1.164	0.908	1.492
Non-Hispanic White	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Lives Alone	<0.001	2.644	2.321	3.011	<0.001	0.518	0.462	0.580	<0.001	0.749	0.648	0.866	<0.001	0.616	0.529	0.718
Lives with Others	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Non-Metro (Participant)	<0.001	0.362	0.300	0.436	<0.001	0.648	0.553	0.759	<0.001	0.609	0.496	0.747	0.006	0.764	0.630	0.926
Metro (Participant)	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
Median Income for ZIP: Particpant Residence	<0.001	0.877	0.838	0.917	<0.001	1.138	1.096	1.182	0.384	0.977	0.928	1.029	<0.001	0.823	0.775	0.874
Percent of ZIP Population: Age 65+	0.008	4.097	1.453	11.555	<0.001	0.118	0.040	0.346	0.894	1.092	0.300	3.981	<0.001	0.036	0.008	0.166
Percent of ZIP Population: Non-Hispanic White	0.006	1.568	1.137	2.163	<0.001	0.567	0.426	0.755	0.478	0.882	0.623	1.248	<0.001	2.525	1.720	3.705

*SC/AAA = Senior Center/Area Agency on Aging; RF = Residential Facility; HO = Healthcare Organization; C/MPC = Community/Multi-purpose Center; FBO = Faith-Based Organization

**Referent Group: Senior Centers / Area Agency on Aging

n = 10,242 (Nagelkerke = 0.243)

Table 3. Participant Personal and Neighborhood Characteristics by Delivery Site Type

Community or Multi-purpose Centers. Compared to workshops attended at senior centers/AAA, participants who attended workshops delivered at community or multi-purpose centers were less likely to be older (OR = 0.970, P < 0.001). These individuals were less likely to have graduated from high school (OR = 0.495, P < 0.001) or have more than a high school education (OR = 0.526, P < 0.001). They were more likely to be Hispanic (OR = 1.356, P = 0.008) or of other/multiple races (OR = 1.666, P < 0.001). Participants who attended workshops delivered at community or multi-purpose centers were less likely to live alone (OR = 0.749, P < 0.001) and reside in rural areas (OR = 0.609, P < 0.001).

Faith-Based Organizations. Compared to workshops attended at senior centers/AAA, participants who attended workshops delivered at faith-based organizations were less likely to be older (OR = 0.977, P < 0.001). These individuals were less likely to have graduated from high school (OR = 0.754, P = 0.008) or have more than a high school education (OR = 0.820, P = 0.047). They were more likely to be African American (OR = 1.360, P = 0.016) or of other/multiple races (OR = 1.725, P < 0.001). Participants who attended workshops delivered at faith-based organizations were less likely to live alone (OR = 0.616, P < 0.001) and reside in rural areas (OR = 0.764, P = 0.006). These participants also resided in areas that were less affluent (OR = 0.823, P < 0.001), had larger proportions of the population who were non-Hispanic white (OR = 2.525, P < 0.001), and had smaller proportions of the population who were age 65 and older (OR = 0.036, P < 0.001).

4. Delivering CDSMP to a diverse set of adults

By examining the Reach and Adoption aspects of this nationwide Chronic Disease Self-Management Program (CDSMP) dissemination, we see the program has the capacity to serve large numbers of heterogeneous older adults through a growing network of delivery site organizations. This initial nationwide roll-out represents many different geographic sites throughout the country and reflects the diversity of older adults' personal (e.g., sex, education levels, race/ethnicity, living situation) and residential characteristics (e.g., residential rurality).

Between the years 2006 and 2009, CDSMP was delivered in 26 states to over 22,000 participants via funding from the Administration on Aging's (AoA) Evidence-Based Disease and Disability Prevention Initiative. In terms of adoption, CDSMP workshops were predominantly attended in areas considered urban (84%), relatively affluent (average ZIP code median income of $56,700 USD), and non-Hispanic white (average ZIP code composition of non-Hispanic white residents was 71%). Overall, workshops delivered at senior centers/AAA (34.8%) and healthcare organizations (26.7%) reached the largest number of participants. In terms of reach, analyses suggest that delivery sites were most successful in recruiting older, non-Hispanic white females residing in urban areas with median incomes of $43,000 USD. However, certain delivery site types were more successful in recruiting diverse seniors, which provides insight into the utilization of delivery site types to attract more diversity among program participants.

To increase the diversity of participants reached by CDSMP, additional efforts are needed to recruit delivery sites that serve diverse populations. For example, to increase CDSMP participation among African Americans, possible strategies may be to encourage more faith-based organizations to offer the program. Or, to increase participation among Hispanic individuals, program adoption among community or multi-purpose centers may be a central focus. Delivering classes at residential facilities and senior centers/AAA may increase program delivery to older participants. Further, more males can be reached by increasing adoption among healthcare organizations and places where older men congregate (e.g., Veteran of Foreign Wars social clubs). And, based on neighborhood-level characteristics, senior centers/AAA and faith-based organizations may be best to reach rural-residing seniors because they are most likely to be adopted in geographically non-metro areas [Table 2] and reach rural-residing seniors [Table 3]. It may also be important to emphasize efforts to recruit participants considered to be at-risk that those attending programs at residential facilities and faith-based organizations resided in less affluent areas.

Given the established effectiveness of CDSMP for improving self-reported health outcomes, it remains especially important to reach seniors more vulnerable to chronic conditions and multiple chronic conditions because these individuals remain at increased risk for premature morbidity and other negative health ramifications. For this reason, it is essential that CDSMP community deliverers build partnerships with those community sectors representing and serving specific populations. Increasing program adoption among a larger and more diverse group of organizations within the aging services network and public health system will improve the delivery infrastructure and enable additional at-risk adults receive these health services. While increasing the delivery infrastructure through partnership building is important, it must be noted that a training infrastructure for community-based Lay Leaders is also needed. Much like recommendations to deliver programs in areas where participants reside, similar considerations should be made when training Lay Leaders to deliver CDSMP in the areas in which they reside (i.e., reducing the time and travel expectations).

This chapter has emphasized Reach and Adoption, initial RE-AIM processes. However, to disseminate self- management programs like CDSMP widely, it is critical to plan for long term-sustainability. In addition to reaching out to partners who can help leverage resources, it is important to develop local program champions who can help sustain and grow evidence-based programs. Similarly, outreach efforts to older adults concerned about managing their chronic conditions can help create demand for these programs. There is a growing evidence-base on the best strategies for reaching out to previously underserved populations. Creating the proper balance between program supply and demand is important and involves continued outreach to potential participants with simultaneous nurturing of Workshop leaders and participating organizations. Another major caution is the need for understanding that this data reflects programs sponsored through the Administration on Aging's initiatives. We recognize there are other CDSMP programs being delivered outside of the aging services network. Unfortunately, there is no systematic data system for collecting participant and delivery setting characteristics.

Current efforts to estimate the cost of evidence-based programs and the cost savings associated with their dissemination are important for strategic planning about how to sustain the program after the end of external grant funding. Guidance for sustainability efforts may be identified by examining processes for building evidence-based programming into existing organizational structures by mandating the use of evidence-based programs for federal funding. Similarly, efforts are underway to examine how lay-led evidence-base programs can receive support from healthcare funding streams traditionally restricted to professional care. Since the roll-out of the 2006-2009 initiative (i.e., the focus of this chapter), there has been increased attention to the potential of evidence-based self-management strategies under the Affordable Care Act (United States Senate). This has enabled CDSMP to be disseminated in 46 States and 2 territories (as of 2010). However, the long-term sustainability of this program is dependent upon a national prevention strategy which makes self-management programs a federal priority that are financially supported (National Prevention Council, 2011).

4.1 Limitations and future directions

This chapter utilized nationally-collected data, which was not without limitation; there are advantages and difficulties when using administrative records for data analysis. First, there were missing data for participants' personal characteristics. This missing data may be attributed to a limited data collection and reporting infrastructure (or fidelity to implemented protocol) in earlier years of the national dissemination. However, trends of missing data were less in later years due to improved data collection. For example, missing data for sex was 24.7% in Year 1, 14.7% in Year 2, and 11.3% in Year 3. Data collection of personal characteristics took place on site and during workshop time. To minimize missing data future program dissemination efforts, it is important to establish data-related expectations for program implementers and reduce the time and resource-related burdens on workshop leaders and program participants. This chapter is limited to findings associated with administrative data. Additionally, this chapter only included data collected from AoA-funded CDSMP workshops, thus it does not represent all CDSMP-related data during the study period. It would be instructive to obtain reports from program deliverers to determine what they identify as the major successes and barriers in the roll-out of evidence-based disease prevention programs for different populations in diverse settings.

Analyses performed in this chapter reinforce the value of using the RE-AIM Framework to assess grand scale translational research efforts/roll-outs in terms of who delivers the program and the types of participants they reach. Although these data provide a valuable glimpse into the reach and adoption-related aspects of this grand scale, national dissemination of CDSMP, less information can be ascertained about the program's implementation, effectiveness, or maintenance. Thus, additional studies are needed to research these aspects. Future efforts to examine the translated CDSMP include monitoring the fidelity associated training procedures and content delivery; determining factors associated with trainer retention, trainers' ability to retain participants, and participant attendance; and examining the influence of distance traveled to workshops (i.e., for participants and trainers) on class attendance and program completion. As ongoing evaluation is critical for continuous quality improvement and seeking external funding, we recommend attention to ways of conducting practical translational research in field settings than can assess self-reported improvement in health-related outcomes among participants (e.g., pain, fatigue, physical functioning) and identify the return on investment or costs associated with healthcare utilization as a result of attending the program (e.g., hospital admissions, emergency room use, preventive screening rates, medication adherence). We foresee the ability to answer some of these questions in a new National Study of Chronic Disease Self-Management which is assessing participants' self-reported health behaviors and health outcomes and linking these outcomes with actual healthcare utilization and cost data (Department of Health and Human Services, 2011).

5. Conclusion

As healthcare resources become increasingly scarce, evidence-based program managers and deliverers at national, state, and local levels continue to explore ways to achieve maximum public health benefit while expending minimal resources. This chapter provides a valuable contribution to these planning efforts by describing how various demographic profiles interact with different community-based program implementation settings. Such knowledge may influence program deliverers to purposively offer programs in settings that reach participants most likely to utilize them. Simultaneously, this chapter provides insight into strategies to increase program adoption in underserved community settings, which have potential to expand participant reach to at-risk and vulnerable senior populations. In addition, this research highlights points of delivery that may be underutilized by practitioners. Demonstrating the value of lesser utilized delivery sites, such as faith-based organizations, is an exceptional way to improve the dissemination of public health programs to improve health outcomes and reduce disparity.

6. Acknowledgement

The national dissemination of Chronic Disease Self-Management was supported as part of the Administration on Aging's (AoA) Evidence-Based Disease and Disability Prevention program (EBDDP) initiative. In addition to the AoA, this initiative was made possible by partnerships with the Centers for Disease Control & Prevention (CDC), Agency for Healthcare Research & Quality (AHRQ), Centers for Medicare and Medicaid Services (CMMS), Health Resources & Services Administration (HRSA), Substance Abuse & Mental Health Services Administration (SAMHSA), and over 30 private foundations. We recognize support from the Administration on Aging with assistance from the National Council on Aging for the evaluation of the EBDDP initiative under cooperative agreement number 90OP0001/01. The findings and conclusions in this article are those of the author(s) and do not necessarily represent the official position of AoA, NCOA, or any other agency. The authors specifically acknowledge the contribution of Nancy Whitelaw, Wendy Zenker, Kelly Horton, Meghan Thompson, Ashley Wilson, Linnae Hutchison, and the CDC Healthy Aging Research Network for their valuable guidance and efforts in the processes leading up to the generation of this book chapter. Additionally, we would like to acknowledge the valuable and continuing work of Kate Lorig and Stanford University with the CDSMP.

7. References

Administration on Aging. (2011). Administration on Aging (AoA) recommendations for grantee quality assurance programs. Retrieved October 1, 2011 from http://www.healthyagingprograms.org/captivate/series1/QA/QA.htm.

Alliance for Health Reform. (2011). *Preventing chronic disease: The new public health*. Washington, DC: Robert Wood Johnson Foundation.

Bolin, J. N., Gamm, L., Vest, J. R., Edwardson, N., & Miller, T. R. (2011). Patient-centered medical homes: Will health care reform provide new options for rural communities and providers? *Family & Community Health, 34*(2), 93-101.

Bookman, A., Harringtonm, M., Pass, L., & Reisner, E. (2007). *Family caregiver handbook: Finding elder care resources in Massachusetts*. Cambridge, MA: Massachusetts Institute of Technology.

Center for Healthy Aging. (n.d.). *Choices for self care: A public-private partnership to improve self care among older adults*. Washington, DC: National Council on Aging.

Centers for Disease Control and Prevention. (2011a). Healthy aging: Helping people to live long and productive lives and enjoy a good quality of life. Retrieved October 2, 2011, from http://www.cdc.gov/chronicdisease/resources/publications/AAG/aging.htm

Centers for Disease Control and Prevention. (2011b). *Sorting through the evidence for Arthritis Self-Management Program and the Chronic Disease Self-Management Program: Executive Summary of ASMP/CDSMP meta-analysis*. Atlanta, GA.

Centers for Disease Control and Prevention (CDC). (2009). Arthritis as a potential barrier to physical activity among adults with heart disease--United States, 2005 and 2007. *Morbidity and Mortality Weekly Report, 58*(7), 165-169.

Department of Health and Human Services. (2011). Request for proposal for a pilot evaluation of the Chronic Disease Self-Management Program: Centers for Medicare and Medicaid Services.

Glasgow, R. E., McKay, H. G., Piette, J. D., & Reynolds, K. D. (2001). The RE-AIM framework for evaluating interventions: What can it tell us about approaches to chronic illness management? *Patient Education and Counseling, 44*, 119-127.

Glasgow, R. E., Vogt, T. M., & Boles, S. M. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM framework. *American Journal of Public Health, 89*, 1322-1327.

Goodman, J. C., Musgrave, G. L., & Herrick, D. M. (2004). *Lives at risk: Single-payer national health insurance around the world*. New York, NY: Rowman and Littlefield Publishers, Inc.

Hughes, S., Seymour, R., & Ory, M. G. (2011). Behavioral risk factors and evidence-based interventions. In T. R. Prohaska, L. A. Anderson & R. H. Binstock (Eds.), *Public Health in an Aging Society*. Baltimore, MD: The Johns Hopkins Press.

Interagency Workgroup on Multiple Chronic Conditions. (2010). *Strategic plan 2010-2015: Optimum health & quality of life for individuals with multiple chronic conditions*. Washington, DC: Department of Health and Human Services.

Lorig, K. R., Bodenheimer, T., Holman, H., & Grumbach, K. (2002). Patient self-management of chronic disease in primary care. *Journal of the American Medical Association (JAMA), 288*(19), 2469-2475.

Lorig, K. R., Ritter, P., Stewart, A. L., Sobel, D. S., Brown, B. W., Jr., Bandura, A., et al. (2001). Chronic disease self-management program: 2-year health status and health care utilization outcomes. *Medical care, 39*(11), 1217-1223.

Lorig, K. R., Sobel, D. S., Ritter, P. L., Laurent, D., & Hobbs, M. (2001). Effect of a self-management program on patients with chronic disease. *Effective Clinical Practice, 4*(6), 256-262.

Lorig, K. R., Sobel, D. S., Stewart, A. L., Brown, B. W., Ritter, P. L., González, V. M., et al. (1999). Evidence suggesting that a chronic disease self-management program can improve health status while reducing utilization and costs: A randomized trial. *Medical Care, 37*(1), 5-14.

Martin, L. G., Freedman, V. A., Schoeni, R. F., & Adreski, P. M. (2010). Trends in disability and related chronic conditions among people ages fifty to sixty-four. *Health Affairs, 29*(4), 725-731.

National Prevention Council. (2011). National Prevention Strategy. Office of the Surgeon General.

Ory, M. G., DeFriese, G. H., & Duncker, A. P. (1998). The nature, extent, and modifiability of self-care behaviors in later life. In M. G. Ory & G. H. DeFriese (Eds.), Self-care in later life: Research, program, and policy issues. New York, NY: Springer Publishing Company.

Ory, M. G., Jordan, P., & Bazzarre, T. (2002). Behavioral Change Consortium: Setting the stage for a new century of health behavior change research. *Health Education Research, 17*(5), 500-511.

Ory, M. G., Liles, C., & Lawler, K. (2009). Building healthy communities for active aging: A national recognition program. *Generations, 33*(4), 82-84.

Ory, M. G., Yuma, P. J., Jarvis, C., & et al. (2006). Doctor-older patient lifestyle discussions in primary care: Analysis of ADEPT videotapes. *Preventive Medicine, 43*(6), 494-497.

Satariano, W. A., Ory, M. G., & Lee, C. (2011). Planned and built environments: Interactions with aging. In T. R. Prohaska, L. A. Anderson & R. H. Binstock (Eds.), *Public Health in an Aging Society*. Baltimore, MD: The Johns Hopkins Press.

Shortell, S. M., Gillies, R., & Wu, F. (2010). United States innovations in healthcare delivery. *Public Health Reviews, 32*, 190-212.

U.S. Department of Health and Human Services. Office of Disease Prevention and Health Promotion. *Healthy People 2020*.
Retrieved October 11, 2011 from HealthyPeople.gov.

United States Department of Agriculture. (2004). Measuring rurality: Rural-Urban Continuum Codes. Retrieved October 1, 2011, from
http://www.ers.usda.gov/briefing/rurality/ruralurbcon

United States Senate. (2009). The Patient Protection and Affordable Care Act as passes section-by section analysis. Retrieved October 1, 2011, from
http://dpc.senate.gov/healthreformbill/healthbill49.pdf

19

Strategy and Practice of Support for Families with Multiple Births Children: Combination of Evidence-Based Public Health (EBPH) and Community-Based Participatory Research (CBPR) Approach

Syuichi Ooki and Kiyomi Hiko
Ishikawa Prefectural Nursing University
Japan

1. Introduction

Since the early 1980s the number of multiple births has rapidly increased in all developed countries largely due to the widespread use of inadequately monitored ovulation induction and multiple embryo transfer (Bryan, 2006). Currently, around 3-4% of live births in Western developed countries are multiple births. Increasing multiple births rates have also been observed in Japan. More than 1% of all maternities (2% of all births) are multiples (Ooki, 2010).

Although there are similarities to singleton pregnancy and parenthood, the experience of expecting and parenting multiples is undeniably very different. The birth and parenting of multiples always present unique challenges for both families and health professionals. Multiple births are associated with substantial medical, health care, socio-emotional, developmental, educational and economic consequences for both families and society (Denton, 2005b; Leonard & Denton, 2006).

Addressing the causes of the high numbers of iatrogenic multiple births continues to be an urgent challenge. Meanwhile, those caring for multiple-birth children and their families need a fuller understanding of their special problems and needs (Bryan, 2006).

In this chapter the author introduces a current public health research and practice integration project for supporting multiple-birth families in Japan, where such support has not yet reached the level of many Western countries. Some important topics related to multiple births are concisely described below.

2. Current support for families with multiples in Japan

Before starting the present project in 2003, we performed a comprehensive book/journal review and information gathering related to the subjects of multiple births and parenting multiples, and we performed key informant interviews to grasp the real situation in Japan. We found very little academic research related to multiples in Japan compared to that in Western countries. Most scientific reports were limited mainly to clinical obstetric researches or twin studies in the field of human genetics. And, few, if any, reports or books related to support for families with multiples were governmental reports or experience-based descriptions or case reports.

2.1 Iatrogenic multiple pregnancy

Multiple pregnancy is the most serious complication of assisted reproductive technology (ART) because of its well established medical risks (Derom et al., 1993; Ooki, 2011c) and social and economic consequences. Many fertile couples in Japan, with or without their knowledge of the health risks and family impact of a multiple birth, are said to regard a twin pregnancy as the ideal outcome of ART. Before any treatment for infertility is performed, the prospective parents need to be fully informed about the risks associated with multiple pregnancy as well as the practical, emotional and financial impact of parenting multiples (Leonard & Denton, 2006).

2.2 Pregnancy and parenting problems related to multiple births

Parents often describe multiple pregnancy as physically and emotionally difficult (Bryan, 2003). The diagnosis of multiple pregnancy frequently comes with a shock as well as happiness. Parents should be given written information about multiple births including details regarding local and national support organizations (Denton, 2005a). The goal of multiple birth education classes for expectant parents is to educate parents and other family members about the unique aspects of multiple pregnancy and parenthood (Leonard & Denton, 2006).

Mothers of multiples have been shown to suffer more from lack of sleep and fatigue than mothers of single-born children. Furthermore, depression is more common well beyond the infancy period, probably due at least to social isolation and fatigue (Bryan, 2003). Child abuse is more common in multiple birth families (Tanimura et al., 1990). The rates of divorce are higher in parents of multiples than of singletons (Jena et al., 2011). The development of most multiple birth children will be within the normal range. However, they will face a higher risk of certain longer-term problems including cerebral palsy, learning difficulties and, in particular, language delay (Bryan, 2003; Denton, 2005b). Multiple birth families have unique needs, which are still not widely understood or sufficiently addressed by health care and other professionals. A well-trained multidisciplinary team, which provides specific care, parent education and support, is the basis for improving health outcomes for multiple birth family members (Leonard & Denton, 2006).

2.3 Financial implications for families

The financial impact of having multiples on families is considerable, with most experiencing a substantial loss in income and an enormous increase in expenditure, especially if the infants are preterm or have complex health needs. Most women leave the work force on reduced or no salary; many do not return to outside employment for months or even years (Hall & Callahan, 2005; Leonard & Denton, 2006). Moreover, extra daily life expenditures are required of multiple birth families compared to families with single-born children. The majority of multiple birth families, despite finding their income greatly diminished, do not qualify for subsidized child care or extra financial support. Early in pregnancy, parents should be encouraged to explore maternal support resources in their families and communities and to seek advice from other multiple birth parents.

2.4 The Declaration of Rights and Statement of Needs of Twins and Higher Order Multiples of ICOMBO

The International Council of Multiple Birth Organizations (ICOMBO ; http://icombo.org/) was formed at the 1980 Congress of International Society for Twin Studies. The mission of the ICOMBO is to promote awareness of the unique needs of multiple-birth infants, children, and adults. The multi-national membership of ICOMBO has developed this Declaration of Rights and Statement of Needs of Twins and Higher Order Multiples as benchmarks by which to evaluate and stimulate the development of resources to meet their unique needs. This statement is a very informative and revealing overview of a variety of socio-psychological and bio-medical aspects of the problems of multiple birth families or multiples themselves. These rights and statements were useful milestones in the development of research and support systems of families with multiples in Japan (Ooki, 2009).

The following statements were adopted in the latest version of the declaration in 2010:

1. Multiples and their families, as any other individuals, have a right to full protection, under the law, and freedom from discrimination of any kind.
2. A. Individuals or couples planning their families and/or seeking infertility treatment have a right to be fully informed about:
    1. factors which influence the conception of multiples;
    2. the associated pregnancy risks and treatments;
    3. the associated risks to one, more or all of the fetuses/infants;
    4. facts regarding parenting multiples; and
    5. the option of multifetal pregnancy reduction along with its associated risks and profound emotional consequences.

   B. Infertility treatment should intend to prevent multiple pregnancies, in particular high order multiples.

   C. Fertility services should disclose their number of multiple pregnancies, both intentional and unintentional.
3. A. Parents have a right to expect accurate recording of placentation, determination of chorionicity and amnionicity via ultrasound, and the diagnosis of zygosity of same sex multiples at birth.

   B. Older, same sex multiples of undetermined zygosity have a right to testing to ascertain their zygosity. Furthermore, involvement in registries of multiple-birth individuals should be absolutely voluntary on the part of the multiples.

   C. Zygosity should be respected as any other human trait and deserves the same privacy rules.
4. Any research incorporating multiples must be subordinated to the informed consent of the multiples and/or their parents and must comply with international codes of ethics governing human experimentation and other types of research.

5. A. Pregnant women, parents and their multiples have a right to care by professionals who are knowledgeable regarding the management of multiple gestation and/or the lifelong unique needs of multiples; and
   B. Multiple births, perinatal and infant deaths, and singleton births that started out as a multiple gestation, must be accurately recorded.
6. Co-multiples have the right to be placed together in foster care, adoptive families, custody agreements, and educational settings.
7. Multiples, as any other human being, have the right to be respected and treated as individuals with their own needs, preferences and dislikes.

The summary of the 'Statement of Needs' is as follows: twins, and higher order multiples have unique conception, gestation and birth processes; health risks; impacts on the family system; developmental environments; and individuation processes. Therefore, in order to insure their optimal development, multiples and their families need access to health care, social services, and education which respect and address their differences from single born children.

2.5 Three aspects of studies on multiples

There are three main independent fields of research regarding multiples, especially twins (Ooki, 2009). Twins are generally recognized as being a valuable resource not only for research on twin births themselves (the study 'of' twins in the field of obstetrics or biology), but also for research clarifying the relative contributions of genetic and environmental factors on human phenotypes (the study 'by' twins in the field of human genetics). It is only more recently that the special problems and needs of twins and higher-order multiple birth children themselves and of their families, and the need for special research (the study 'for' twins in the field of maternal and child care in public health) have begun to be recognized, although the number of reports is still very small compared to the other two research fields.

"Scientists must be encouraged to investigate norms for developmental processes which are affected by multiple births, benchmarks of healthy psychological development, and strategies and interventions that are effective in promoting the health of multiple birth families during the parenting period" (ICOMBO).

These three types of studies on twins and multiples as essentially strongly related to one another. The family support practices of families with multiples will become more fruitful if the findings of scientific research concerning twins and multiples in biology, obstetrics, psychology and human genetics are taken into account, in addition to the research on maternal and child care. The present project put strong emphasis on this point.

2.6 Support for families with multiples in Japan

The Japanese Association of Twins' Mothers (JATM) was established in 1967. Although JATM has the longest history of activity and the largest number of members throughout Japan, a nationwide systematic support system for families with multiples has not yet been achieved. There are many small local twins' mothers' clubs (so-called Mutual-Aid Self-Help groups) throughout Japan, although the exact number of such groups is unknown. The reason for this is that they come into existence and disappear in a relatively short time. Most of these small clubs do not have a stable foundation in terms of both financial and human resources. Many clubs have had common problems, for example the lack of a successors to a club leadership, relatively short periods of enrollment of the members, etc. There are no organizations connecting these small clubs (Ooki, 2009).

Many Western countries have nationwide organizations to support multiple birth families (Denton, 2005a). Support from local governments, public health centers and municipal maternal and child health care centers are not sufficient in Japan (Ooki, 2009). The history of such support is very short; e.g., the first multiple birth families-focused child care class was held in 1991. Very little information, advice and support is provided by medical institutions. Even if a maternity class regarding multiple pregnancy is held, most of these activities are not followed up with other similar programs in the communities where families live once the mothers deliver the multiples and leave the hospital. Multiples-focused pregnancy and parent education resources include leaflets, books and other printed materials, videos and DVDs, prenatal childbirth education classes, and online or telephone information and support networks. Valuable peer and professional support may come from local multiples' support groups or health care centers (Leonard & Denton, 2006). Although the situation has gradually improved, there remain very few such resources in Japan.

3. Public health challenges related to multiple births in Japan

All of the problems related to multiple pregnancy cannot be resolved if they are discussed merely in the context of certain families having problems at the moment or as problems in the narrow sense of clinical obstetrics. It is very difficult for public health professionals to recognize that families with multiples need special support, since multiple births themselves are not a disease or impairment. The main public health challenges related to multiple-births families in Japan can be summarized as follows (Ooki, 2009).

3.1 Lack, division, disparity and inequality of information

All kinds of information regarding multiple births is lacking, including evidence-based objective data on, for example, fertility treatment and short- and long-term prognosis of multiples. Qualitative data, for example, tips on the child-rearing of multiples, are also lacking and not fully described and organized. Therefore, health professionals do not know what the problems are. Moreover, there is very little cooperation and collaboration between families with multiples, medical institutions, administrative agencies, educational institutions and research institutions. This situation leads to division and inequality of information between parties and within each party.

3.2 Delay of institution of laws and guidelines

Governments and academic societies have very few guidelines or policies regarding multiple births, including fertility treatment, compared to Western countries or national or international academic societies. Therefore, the future directions of support for multiple birth families are vague and unclear.

3.3 Lack of specialists in multiple births

In general, few professional advisors in the fields of pregnancy and the growth and development of children have all kinds of adequate information to answer parents' questions concerning parenting multiple birth children.

3.4 Shortage of infrastructure and social resources

Social supports for families with multiples are very limited, and not necessarily convenient to use. Moreover, these social resources are very different among municipalities in both quantity and quality.

3.5 Lack of mental support for multiple families

Since multiple pregnancy has high risks, the priority is put on the safety of the delivery, and as a consequence mental care for mothers is overlooked. It is also important to improve family competency, including the capacity to make informed decisions, and to empower families.

4. Conceptual model of the present project

These problems surrounding multiple birth families are never resolved by the efforts of families with multiples alone, even if they create local twins' clubs or groups. A more multidisciplinary collaboration including specialists from each domain involved is essential. Moreover, population-based or at least large-scale epidemiologic studies to assess the long-term health and the social and psychological impact of multiple births on family, children and society are crucial to provide a scientific basis and to persuade policymakers of the importance of supporting families with multiples. The key concept of present practice is outlined in Figure 1. We are trying to combine research and practice by applying the EBPH (Evidence-Based Public Health) method (Gray, 2008), the CBPR (Community-Based Participatory Research) approach (Israel et al., 2008) and a population approach (Rose, 1994). Our emphasis is on determining the real needs of community members.

4.1 Needs of a population approach and a high-risk approach

The level of the health conditions of multiple birth families is lower than that of families with single-born children. There is no maternal and child health policy focused on multiple birth families in Japan. These disadvantages and this inequality of information start during pregnancy, or even during fertility treatment. This situation raises the risks of anxiety, experiencing difficulty and of failure in the child care or parenting of multiples. It is obvious that ad hoc individual support without a clear purpose, as is presently available, is insufficient, and a population approach focusing on all multiple birth families would be more effective, as shown in Figure 2. Merely waiting for families to ask for help or to take part in the child care classes is less effective than providing information actively and positively to all multiple birth families. It is important to target families without information, families who have information but cannot utilize it, as well as families who have information and can use it, as shown in Figure 3.

On the other hand, multiple birth families are obviously a high-risk group in terms of maternal and child health. Moreover, many future risks are predictable during pregnancy or fertility treatment. Appropriate early intervention as a 'high-risk approach' would also be effective. Community networks for multiple birth families would make both the population approach and the high-risk approach possible.

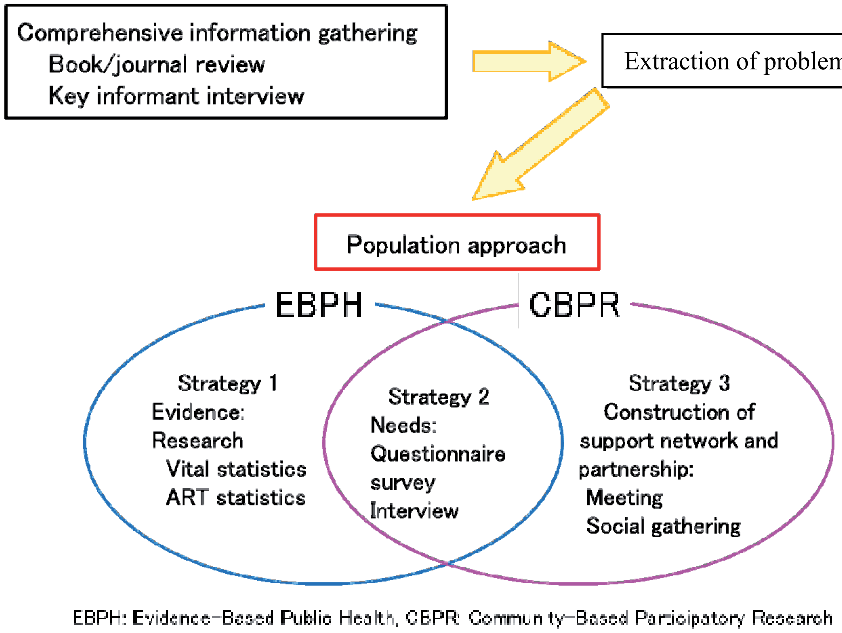


Fig. 1. Conceptual framework of the project.

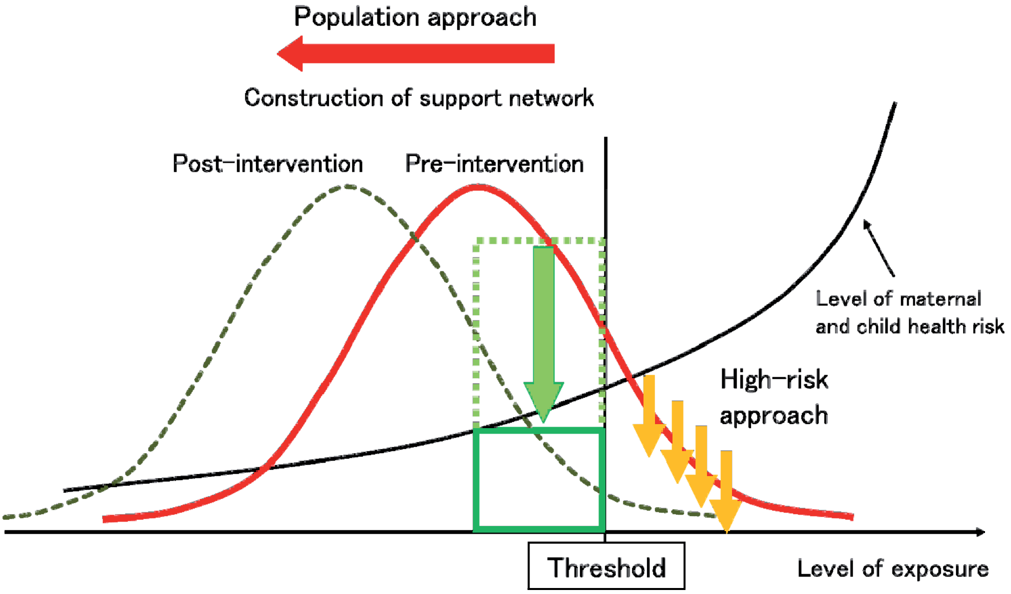


Fig. 2. Construction of support network for families with multiples as a population approach.

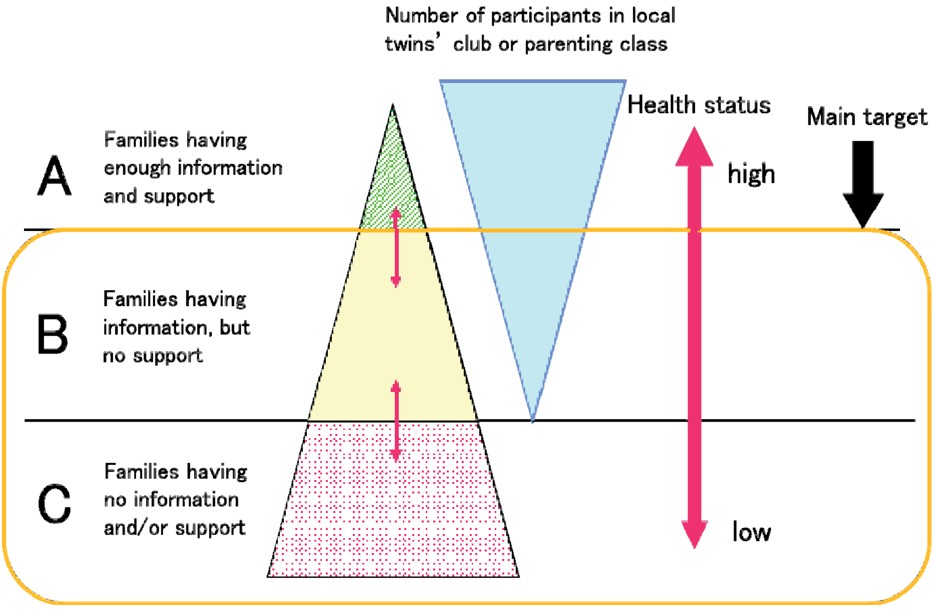


Fig. 3. Health status of families with multiples in the community.

4.2 Evidence-Based Public Healh (EBPH)

As shown in Figure 4, the three main components of Evidence-Based Public Health (EBPH) are evidence, value and local needs and resources (Gray, 2008). EBPH requires that decisions about public health practice are based on the best available, current, valid and relevant evidence. These decisions should be made by those receiving care, informed by the tacit and explicit knowledge of those providing care, within the context of available resources. EBPH can also be defined in population health terms, e.g., "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of communities and populations in the domain of health protection, disease prevention, health maintenance and improvement (health promotion)". Evidence can come from a range of sources including scientific journals and other publications, population health statistics, epidemiologic data and locally collected data.

4.3 Community-Based Participatory Research (CBPR)

CBPR is an action research approach that emphasizes collaborative partnerships between community members, community organizations, health care providers, and researchers to generate knowledge and solve local problems (Israel et al., 2008). Hierarchical differences that typically arise between providers and patients are flattened through this partnership and all participants work together to co-create knowledge and effect change throughout all aspects of the research process. Each partner contributes unique strengths and knowledge to improve the health and well-being of community members. CBPR has gained increased credibility in healthcare and public health since the early 1990s because of its potential to inform understanding of individuals' health related experiences and inform the creation of workable and appropriate services. As shown in Figure 5, first, the researchers must know what issues the community members are willing to work on, and they must understand the related attitudes, values, beliefs and customs ('Felt needs').

Fig. 4. Three main components of EBPH (Evidence-Based Public Health).

Fig. 5. Community partnership: Interlocking Dynamics model in the CBPR (Community-Based Participatory Research) approach. (Permission from Noel J Chrisman (1997). UW, personal communication)

The following nine principles seek to capture key elements of the CBPR approach (Israel et al., 2008). 1. CBPR recognizes community as a unit of identity. 2. CBPR builds on strengths and resources within the community. 3. CBPR facilitates collaborative, equitable partnership in all research phases and involves an empowering and power-sharing process that attends to social inequalities. 4. CBPR promotes co-learning and capacity building among all partners. 5. CBPR integrates and achieves a balance between research and action for the mutual benefit of all partners. 6. CBPR emphasizes public health problems of local relevance and also ecological perspectives that recognize and attend to the multiple determinants of health and disease. 7. CBPR involves systems development through a cyclical and iterative process. 8. CBPR disseminates findings and knowledge gained to all partners and involves all partners in the dissemination process. 9. CBPR requires a long-term process and commitment to sustainability.

4.4 Three strategies in the present project

The first strategy is to monitor secular trends of vital statistics and ART statistics concerning multiple births, although the latter are very limited in Japan. Monitoring and reanalyzing these statistics provide an important objective macroscopic view of the public health problems related to multiple births.

The second strategy is to provide evidence-based information to health professionals and policymakers as well as multiple birth families. A large-scale database of multiples, mainly twins, was begun in 1987 (Ooki & Asaka, 2005, 2006). The primary purpose of this database was genetic epidemiologic twin research. However, it turned out that the data, including both quantitative and qualitative, were also a useful resource for the provision of information on many features specific to twins, for example their growth and development, to families with multiples and health professionals. There are many tips for parenting multiples which are derived mainly from experience and intuition. Scientific data will add correct evidence-based information to these experience-based tips.

The third and most important strategy is to construct a human network and family support systems at the prefectural level by the CBPR approach, as shown in Figure 6. It would be very difficult to construct a nationwide family support network quickly. Therefore, prefectural level support networks should be constructed initially. Although the author may focus mainly on twin pregnancy and parenthood in the following description, the demands associated with higher-order multiples are of course similar but usually greater.

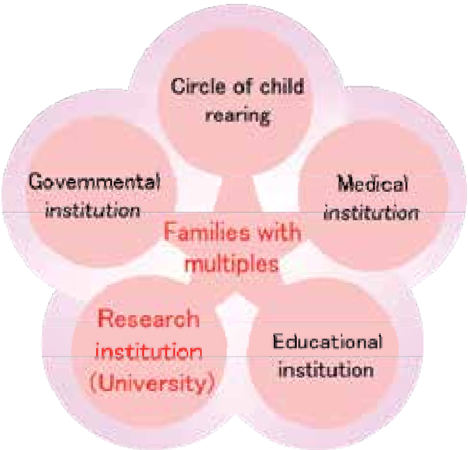


Fig. 6. Continuous and comprehensive support through pregnancy, child births and child rearing in a community support network.

5. First strategy: Making good use of vital statistics and ART statistics

All available vital statistics on multiple births within the entire Japanese population since 1974, originally collected by the Ministry of Health, Labor and Welfare (MHLW), were gathered, combined and reanalyzed. The vital statistics are based on birth records, which is published as an annual report of aggregate, not individual, data.

5.1 Secular trends of multiple births rates

Recent multiple births rates have been affected mainly by fertility treatments and maternal age (Derom et al., 1993). Secular trends of multiple births rates increased twice during the past two decades, and about 2% of all neonates are now multiples (Ooki, 2010). About 20 thousands live births are multiples, as shown in Figure 7.

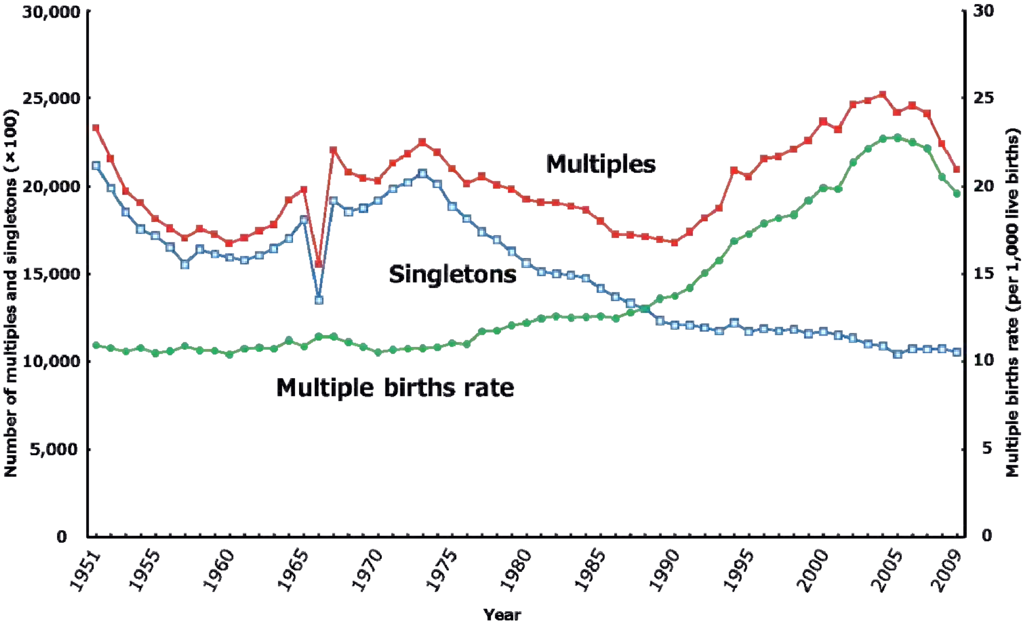


Fig. 7. Secular trend in the number of live birth singletons and multiples and the multiple births rates.

5.2 Low birthweight and preterm delivery

The population attributable risk percent was a useful indicator for clarifying the public health impact of certain risk factors. The population attributable risk percent of both low birthweight and preterm delivery tended to increase as a whole during the last thirty years, and recently reached around 20%. This is in relatively good accordance with several international studies (Blondel et al., 2002, 2006). The risks of very preterm delivery (<32 weeks) and extremely low birthweight (<1,000g) attributable to multiples are much higher than the risks of overall preterm delivery (<37 weeks) or low birthweight (<2,500g). Very preterm delivery or low birthweight require intensive care in neonatal units, and are at high risk for neonatal morbidity and developmental problems. Therefore, the rising number of multiples will increase the burden on neonatal services and health services in general, as well as resulting in higher numbers of children surviving with impairment.

5.3 Stillbirths rates, perinatal and infant mortality

Inasmuch as all these indicators were markedly improved both singletons and multiples, multiples were still 2-5 times higher compared with singleton (Ooki, 2009). The main reason for this elevated perinatal mortality is preterm and very preterm birth, resulting in low and very low birthweight children. But, it is worth noting that prognosis is rather good as to multiples than singletons, if certain period of gestational weeks or birthweight is obtained. According to vital statistics, the infant mortality rate is lower in multiples compared to singletons, if birthweight or gestational age is near the cut-off point of LBW (2,500g) or preterm delivery (37 weeks).

5.4 Composition of maternal age

Maternal age class of multiple births is higher than that of singletons (Ooki, 2011a). About 60% of mothers were estimated to have given birth to their multiples in the first delivery between 1999 and 2009. It seems that mothers of higher age have many physical and mental burdens for parenting multiples. Moreover, if mothers take fertility treatment, economic burden would be also added. This may be still the case as to fathers.

5.5 The impact of maternal age and fertility treatment

During the 25-year period, multiple-birth rates according to maternal age class increased after the late 1980s. This tendency was obvious in women aged 35 to 39 years (Ooki, 2011a). The estimated numbers of iatrogenic multiple births greatly increased in women aged 30 to 34 and 35 to 39 years. The rate (per 1,000 live births) of iatrogenic multiple births gradually increased from 0.7 (1977) to 1.3 (1986), then rapidly and markedly increased from 1.3 (1986) to 11.4 (2005), and finally decreased to 8.1 (2009). The estimated maximum percentage of iatrogenic multiple births was 50.0%, in 2004 and 2005. The rapid increases in Japan in the number and rate of multiples born to women older than 30 years are likely due to iatrogenic rather than spontaneous multiple births, as shown in Figure 8.

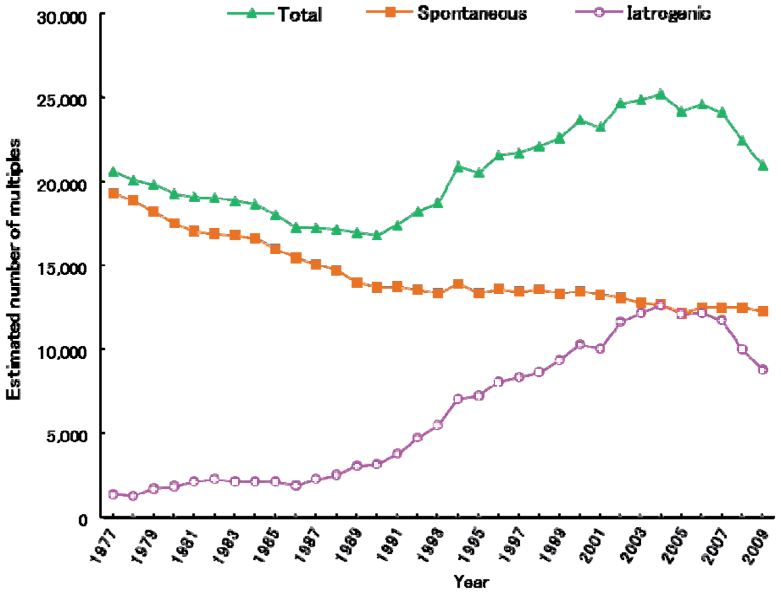


Fig. 8. Secular trends in the sum of the age-specific estimated number of spontaneous and iatrogenic multiple births, 1977-2009.

The effect of ART and non-ART ovulation stimulation fertility treatment on the number and rate of multiple live births from 1979-2008 was estimated using two independent data sources, ART statistics presented by the Japan Society of Obstetrics and Gynecology between 1989 and 2008 and vital statistics (Ooki, 2011b). The number of non-ART iatrogenic multiple births was estimated by subtracting the ART multiples from the total iatrogenic multiples. There was an overall increase in the non-ART multiple births during the thirty year period, whereas ART multiples tended to increase from 1983 to 2005, and then rapidly decreased thereafter, suggesting the effect of the single embryo transfer policy for ART. The number or percentage of ART multiples was almost consistently lower than that of non-ART multiples.

5.6 Summary

The present results offer clear evidence of the public health impact caused by the rapid increase of multiple births in Japan. Caution is recommended when considering these results, however, because these indicators were counted by the number of children (multiples), not by the number of mothers (families). According to the vital statistics in Japan, the mean number of children under 6 years old in one family was larger in families with multiples compared to families with singletons. Therefore, the risk of having stillbirth and/or child death in at least one baby in one family may be even higher in families with multiples than in families with singletons. Moreover, these impact can be presented by medical economics (Hall & Callahan, 2005), laws and guidelines on fertility treatment or multiple births (Bryan et al., 1997; Denton, 2005b; Leonard & Denton, 2006), the real situation of families with multiples (Thorpe et al., 1991), and social family support system or maternal and child health policies (Ooki, 2006b).

6. Second strategy: Construction of large-scale twin database and provision of evidence-based information

Many countries including some Asian countries are constructing or have constructed large population-based twin registries, mainly for genetic research (Ooki, 2009). No systematic twin registry exists in Japan, however. Since the frequency of families with multiples is not so high (1% at most), a strategic method is crucial, if researchers hope to gather a large amount of high-quality unbiased data.

6.1 Strategies for the collection of data on multiples in Japan

There are four main types of data included in studies on multiples (Ooki, 2006b). First, vital statistics can be obtained (Kato, 2004), but it is almost impossible to obtain access to personal information concerning individuals. Second, data from large hospitals have been used in the field of obstetrics. The collection of obstetric data on multiples is relatively easy with the trade-off of selection bias in favor of high-risk infants. Third, the Basic Resident Registration of municipalities can be used. But, cost-effectiveness is extremely low in the case of multiple birth families. Fourth, there is a volunteer-based database of multiples, which includes data from mothers belonging to associations for parents of multiples. It contains more detailed information on the condition of multiples after birth compared to the vital statistics and hospital data. Although volunteer-based databases may have some selection biases, cost-effectiveness is very high.

A volunteer-based twin database has been organized from 1987 (Ooki & Asaka, 2005, 2006) that is larger and less biased than hospital data and contains more detailed information after birth. Collaborative research with multiple-births families using the CBPR approach could increase the number of participants, improve the response rate for the questionnaire survey and validity of the data, provide researchers with information about the real needs of families and assist researchers in creating health needs-oriented research questions. In the CBPR approach, researchers become participant observers, as shown in Figure 9.

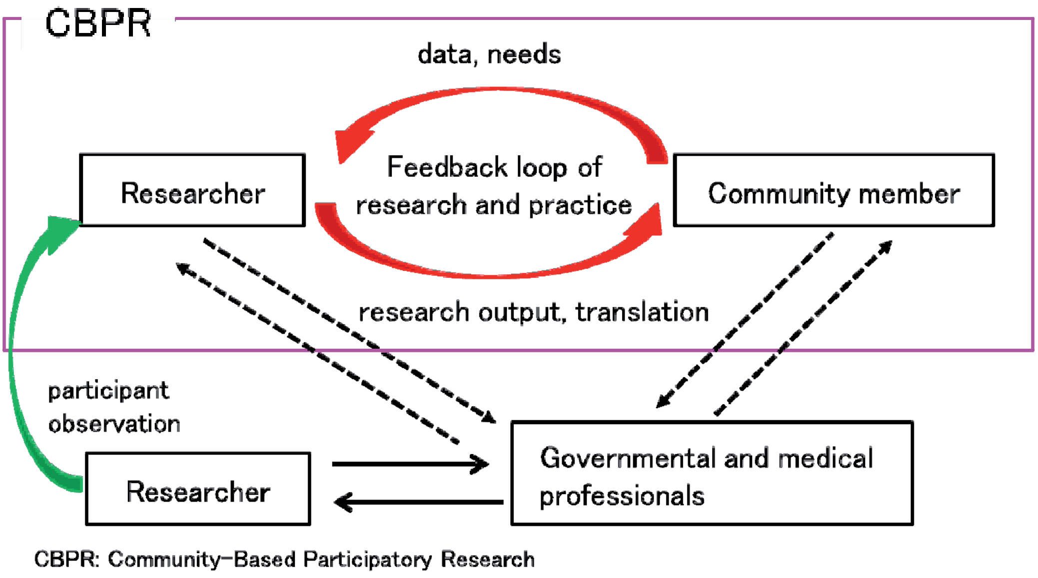


Fig. 9. Collaborative research with community members based on partnership.

6.2 Outline of Japanese Database of Multiples in Childhood (JDMC)

The database consisted mainly of three independent groups. The first group -school applicants group- included 1,205 mothers and their twin children, who had applied between 1981 and 2011 to the secondary school attached to the Faculty of Education at the University of Tokyo (Ooki & Asaka, 2005, 2006). The second group – the maternal associations group – consisted of 951 mothers from several associations for parents of multiples throughout Japan. Continuous data have been gathered from 2001(Ooki & Asaka, 2005). The third and most recently formed group consisted of 956 mothers who were recruited during collaborative research with multiple-births families throughout Japan starting in 2011, based on the present CBPR approach. Mailed or hand-delivered questionnaires were used to collect the basic data. Most medical findings of all databases were obtained from *The Maternal and Child Health Handbook*, which is presented by the MHLW to all pregnant women. This handbook serves as a valuable source of health information for mothers and children up to 6 years old. These questionnaire surveys are now in progress.

6.3 Ethical issues

As to school applicants group, the statistical analysis of the data was clearly written in the application document, and the detailed explanations concerning data collection by questionnaire and interview, and blood sampling for zygosity examination and health check were added as another paper from 1999. Moreover, informed consent was obtained from each twin and his or her parents in writing from 2001 on. The data analysis was also permitted by the ethical committee of this school. Zygosity diagnosis using DNA sample was permitted through the ethical committee of the Graduate School of Medicine, University of Tokyo. All the mothers in the maternal associations group cooperated voluntarily in this research, mainly through the presidents of their associations.

6.4 Zygosity diagnosis

6.4.1 Method of zygosity determination

There are two types of twins. Monozygotic (MZ) twins derive from the division of a single zygote, whereas dizygotic (DZ) twins derive from the independent release and subsequent fertilization of two ova (Machin, 1994). Zygosity determination is the process of determining whether same-sex twin pairs are MZ or DZ. The three reasons for determining zygosity at birth are (1) medical, (2) scientific and (3) personal (Derom et al., 2001). Zygosity is particular importance in questions of organ transplantation and inheritance of specific diseases (medical). To determine the relative contribution of genetic and environmental factors to the human traits, comparison of similarity between MZ and DZ pairs is made in the field of human genetics (scientific). The question of zygosity has a special importance to the multiples and their families (personal). The needs for an appropriate method of determining zygosity for use by twins' parents or health professionals have increased. Although the accuracy was lower concerning questionnaire method compared to DNA markers, cost performance is high and less invasive.

6.4.2 Development of zygosity questionnaire for Japanese twins

The determination of zygosity in twins based on questionnaires can be done with considerable accuracy, showing that the accuracy of the questionnaires employed is around 95% (Rietveld et al., 2000). The zygosity questionnaire for twins' mothers by Ooki et al. (1993), consisting of five simple questions regarding twins’ similarity, was upgraded and achieved a total accuracy of 94.6% (212/224 pairs) (Ooki & Asaka, 2004). If zygosity determination by DNA/genetic markers was regarded as the gold standard, the accuracy of the zygosity questionnaire was 97.5%, leaving around 10% of pairs unclassified. The effectiveness of this simple method in practical use, especially when this questionnaire is used for the purpose of offering zygosity information to twins' mothers and health professionals more easily.

6.5 Representativeness of the subjects of JDMC

Most of the present subjects were normally developed twins. Therefore, birthweight seemed to be larger than the general twin population. Compared with hospital data, however, the present data more closely reflects real birthweight of the general twin population. It was ascertained that there was no fatal selection bias of severely discordant twin pairs. But some unexpected selection biases may exist (Ooki & Asaka, 2005).

6.6 Main evidences obtained from the JDMC

Main evidence produced by the present database (Ooki, 2009), which was provided to the multiple birth families and health professionals is described in brief.

6.6.1 Zygosity misclassification of twins at birth

By analyzing four independent samples, we found that at a constant rate about 25%-30% of MZ twins were misclassified as DZ twins at birth (Ooki et al., 2004). This percentage is in very good accordance with that of MZ twins having dizygous placenta. About 70% of mothers showed their interest in their children's zygosity for many reasons. One of the main reasons was their doubt about the zygosity they were told at birth. Considering all of these issues, medical specialists should be careful what advice they give parents about zygosity, placenta and related information.

6.6.2 Growth and development of twins

In Japan, the developmental norms have been examined every ten years and have been presented in *The Maternal and Child Health Handbook* by the MHLW. The physical growth after birth, motor and language development of twins in childhood must be evaluated using the standards for the general population, that is, essentially the standards for singletons. According to these measurements, many twins are regarded as having poor growth and development, especially when they are very young, and this causes both the twins and their parents much embarrassment and concern. Appropriate data should be provided to families and health professionals.

6.6.2.1 Intrauterine growth of twins

Because birthweight is the strongest indicator of the risk of perinatal death, birthweight norms are important both for clinical practices and for epidemiologic studies. Numerous birth weight standards by gestational age for twins have been reported in Western countries (Glinianaia et al., 2000), whereas little is in Japan (Kato, 2004). Other body size parameters at birth, such as birth length, chest circumference, and head circumference (Ooki & Asaka, 1993) have not been reported as consistently as birthweight. Selected percentiles by gestational age were calculated and smoothed (Ooki & Yokoyama, 2003). Compared with singletons, birth weight difference in twins was marked and slight difference was observed as to length, whereas no difference was observed as to chest and head circumference. It is very important for both multiple birth families and health professionals to recognize that the intrauterine growth of twins is amazingly different from that of singletons.

6.6.2.2 Physical growth of twins in childhood

There have been many studies on the physical growth of twins in childhood in Western countries (Akerman & Fischbein, 1992; Wilson, 1986). Most studies reported so far regarding the physical growth of twins after birth in Japan had very small samples and only roughly classified age after birth. Using 2,029 pairs of normally developed Japanese twins, Ooki and Yokoyama (2004) analyzed the characteristics of twins' physical growth from birth to 6 years and presented growth charts based on the analysis. The size deficit of the twins compared to the standards for the general population of Japan was largest at birth: more than 20% for weight and approximately 6% for length. These deficits decreased rapidly in the first 6 to 12 months, and were found to be as low as 0-2% at 4 to 6 years of age. Growth charts specifically for twins are needed, at least for the first 1 to 3 years of age but not beyond the age of 6 years.

6.6.2.3 Longitudinal similarity of body weight from birth to 3 years

To clarify longitudinal similarity of body weight, 648 pairs of same-sex twin pairs were analyzed (Ooki, 2009). More than 90% of MZ pairs were similar, whereas 65% as to DZ. About half of MZ pairs were constantly similar after birth. Even if individual twin baby are born small, they rapidly catch up with singletons until one year of age. MZ pairs become more and more similar by one year of age irrespective of birthweight difference, whereas DZ pairs become more and more dissimilar. This tendency becomes marked with age. This is because MZ twin pairs, who have identical genetic composition, reach to the genetically determined their body weight. On the contrary, DZ twin pairs, whose genetic similarity is only 50%, equal to siblings in general, show variety of similarity.

6.6.2.4 Motor development of twins

Many studies have shown that twins tend to lag behind singletons in terms of motor development (Chaudhari et al., 1997; Peter et al., 1999), although the causes of this delay are not explained simply by the earlier gestational age at birth or the lighter birthweight of twins compared to singletons. Relative to general population norms in Japan, twins tended to be delayed in reaching several motor development milestones (Ooki, 2006a). Nevertheless, as twins mature, their development tends to catch up with that of singletons. Stepwise regression analysis showed that gestational age was the most influential on all six milestones. Motor development within pairs was more similar in monozygotic pairs compared with dizygotic pairs regarding each item, suggesting genetic contributions. More than half of the DZ pairs showed differences of one month or more in attaining walking without support.

6.6.2.5 Language development: Age at first spoken word

It has become a well-known fact that twins tend to lag behind singletons in terms of language development (Rutter et al., 2003; Thorpe et al., 2003), although the causes of this delay are unknown and appear to be complex. The subjects were 937 (maternal associations group) and 1,092 (school applicants group) twin mothers and their twin children (Ooki, 2005). Relative to general population norms, twins in the maternal associations group were about one to two months delayed in terms of age at first spoken word. Nevertheless, about 95% of twins have started speaking at least one word at 18 to 19 months, as singletons do. According to maternal reports, MZ males, compared with singletons, showed the largest delay throughout childhood; this delay was partly attributed to the effect of 'twin language'. Language development was more similar in MZ pairs than that of DZ pairs. The language development of twins partly differs from that of singletons and that estimation of language development specifically for twins are needed.

6.6.2.6 Breast-feeding of twins

Although an adequate quantity and quality of milk production has been documented even for high multiples, it seems difficult for mothers rearing multiples to breast-feed for many reasons. Many studies deal with the introduction of the skills or techniques for breastfeeding multiples (Flidel-Rimon & Shinwell, 2006; Gromada & Spangler, 1998), and practical recommendations or guidelines (Leonard, 2003). Breastfeeding rates from 0 to 6 full months were analyzed using 4,023 Japanese twins (Ooki, 2008). Full breastfeeding of twins has risen recently in Japan, although the rates are still lower than those of the general population. But, the combined rates of full and partial breastfeeding were close to those of general population. The most influential factor that interfered with breastfeeding from 0-6 months was gestational weeks. The percentage of concordance pares were around 95% for all months.

6.7 Some practical guidelines

The results of scientific research should be offered not only to researchers and professionals of maternal and child health, but also to families who offered valuable data (participants). The reliance between researchers and research participants or collaborators (families with multiples) is crucial, as shown in Figure 9. Research results must be translated into information that is useful for family support activity. There are two essentially important and often confused dimensions to estimating the growth and development of twins: comparing individual twins versus singletons, and comparing the similarity of the twins within a pair.

6.7.1 Physical growth and motor and language development

Twins showed a tendency to lag behind singletons in terms of physical growth and motor and language development in the present successive studies, as many studies have already pointed out. Nevertheless, as twins mature, their development tends to catch up with that of singletons. According to the interviews with mothers, the following advice was frequently given upon health examinations by health professionals without decisive evidence. *'Because your children are twins, delay of the growth and development compared with singletons, is a natural event'*, or *'Your children will catch up with singletons sooner or later. Therefore, do not worry'*. Surely this advice offers temporary comfort to mothers, but provides no essential solution. Concern about growth and development, especially language development, is one of the most common and most serious questions about their children that parents, especially those of twins, bring to their pediatricians or public health nurses. Although a substantial proportion of delays in twins seemed to resolve themselves in the preschool or early school years. Moreover, MZ pairs are in general more similar than DZ pairs, with variety of degrees, concerning almost all traits of growth, development and habitual behaviors. In other words, DZ pairs show considerable within-pair differences or discordance. Therefore, it is of little meaning when estimating or advising the similarity of twin pairs without considering their zygosity (Ooki, 2009).

Since about two-thirds of all twin pairs in Japan are DZ pairs due to widespread fertility treatments, dissimilarity within pairs is usually seen. The professional caregiver should abandon the stereotypical notion that all twins share similar characteristics, which is one of the most popular and erroneous ideas about twins in Japan. Similarity of twins were strongly influenced by zygosity. Zygosity is one of the key concepts connecting genetic studies with the maternal and child health care of twins. In the situation of a medical examination of twins, health professionals should be aware that even within twin pairs, growth and development can differ considerably, and the existence of difference or discordance within a pair is usually in itself of little consequence. Moreover, total estimation of growth and development as twin children is essential, which should be take resemblance and difference of twins from singletons into consideration.

6.7.2 Breast-feeding multiples

Combined rates of full and partial breast-feeding were close to those of the general population. It was desirable to raise the full breast-feeding rates of multiples while maintaining the total (full and partial) breast-feeding rates. Antenatal counseling, hospital practice, the attitude of the medical team towards breast-feeding, the expertise of the public health nurse, and national policies on maternal and child health could change the current situation. However, when breast-feeding is not possible, the health professionals must carefully avoid judgmental approaches that may induce guilt, and mothers should not be given the impression that they have to breast-feed exclusively in order to breast-feed successfully (Leonard, 2003).

7. Third strategy: Human network construction in Ishikawa Prefecture

7.1 Background

The majority of twin registries throughout the world have been constructed primarily for genetic studies. It appears to be very difficult to achieve a high participation rate from families with very young children in Japan, particularly if researchers perform only a genetic twin study with no feedback for the participants. The nurturing of multiples entails a higher burden physically, mentally and economically than that of singletons, and participants surely expect appropriate information from researchers to facilitate the healthy growth of their twins. Offering information useful for the parenting multiples would be a strong incentive for the parents to participate in such studies, as shown in Figure 9.

Given this background, the construction of a population-based database of multiples in childhood at the prefectural level began in 2004 (Ooki, 2006b). The basic idea is a population approach to reduce health risks of multiple birth families. The goals of the registry are to contribute to the development of welfare programs for multiple birth families as well as to co-ordinate research useful for both human genetics and maternal and child health. The well established and sophisticated strategies and methods to recruit twin families into the registry were very useful for the construction of human network and information distribution in this program.

7.2 Maternal and Child Health Administration in Japan

Japan consists of 47 prefectures, the basic unit of local government, and about 1800 (1719 as of 2011) municipalities. National government policies for the health of mothers and children are planned and administered by the MHLW. Public Health Centers (PHCs) were operating in approximately 517 locations in 2008 (495 as of 2011); additionally, the government offices of cities, towns and villages operate other municipal PHCs. These health centers administer independent policies together with the policies and administrative functions delegated or transferred by the MHLW. At present, most of the functions of Maternal and Child Health administration have been transferred from the prefectural level to the municipal level. Typically, a single prefecture has several PHCs, which serve several municipalities within their catchment area. PHCs establish communication and coordination between municipalities with respect to maternal and child health projects in cities, towns and villages, give guidance and advice to municipalities on technical matters, and provide expert maternal and child health services. In other cases, the city defined by law as the 'core city' of the prefecture has its own PHC.

7.3 Ishikawa Prefecture

This Prefecture is located in the middle of the Hokuriku region of Honshu Island. The former feudalism may be reflected in attitudes toward patriarchy, sterility, child nurturing and multiple births in some districts. Kanazawa city, the prefectural capital, is now the center of the Hokuriku area. When this project was initiated in 2004, the total number of municipalities was 39 (19 as of 2011). The birthrates of this prefecture over the past 10 years have been nearly the same as those of Japan as a whole. Ishikawa Prefecture has four prefectural PHCs, each with a branch office, and Kanazawa City, the core city, has three PHCs.

7.3.1 Vital statistics about the multiple birth rates

To obtain an initial outline of total multiple births, secular trends of multiple births in Ishikawa Prefecture were analyzed based on the Japanese vital statistics. The twin birthrate of this prefecture is higher than that of Japan as a whole. The number of multiple births is between about 100 and 160 deliveries each year between 1995 and 2008 (91 as of 2009). This number makes exhaustive identification of newborn multiples and a construction of the population-based approach possible.

7.3.2 Social support by governmental and medical institutions

Sources of support for multiple birth families and information provided as support by governmental and medical institutions were compiled exhaustively from a mailed questionnaire in June 2004 (Ooki, 2006b). Recipients were all PHCs in the prefecture, a municipal PHC, and obstetric and pediatric medical institutions. The number of surveyed institutions totaled 417. The number of support associations available to the parents of multiples was ascertained through governmental and medical institutions as well as through personally obtained information, and their activity was investigated. The three most important problem areas according to governmental and medical professionals in supporting multiple birth families were the following: 1. the lack of knowledge on multiples among the health professionals themselves, 2. the lack of information on the multiple birth families or on the multiple birth itself, and 3. insufficient social resources available to multiple birth families. Multiple births deliveries were highly concentrated in the district at several higher-level medical institutions with neonatal intensive care unit. There are 11 associations for parents of multiples, six in Kanazawa. The present findings and outline of social support organizations available to families with multiples were important in the development of a closely focused case report and in the construction of an effective human network.

7.4 Construction of the human network

A human network to support multiple birth families was organized alongside the demographic research and questionnaire survey using the CBPR approach. This network was constructed with the help of the relationships between families with multiples, support groups for child rearing, governmental and medical institutions, and university. The Health and Welfare Bureau of Ishikawa Prefecture provided assistance in a positive way: several intensive meetings were held for the purpose of exchanging information between members of associations for the parents of multiples, the medical staff, public health nurses, midwives and twin researchers. Workshops, round-table discussion and other events were also held periodically, about two or three times per year, in both the local and central districts. The workshop program included professional lectures on multiples and meetings where the parents of multiples could meet with each other and exchange experiences. Information of this network was vigorously presented to the mass media, including television stations, newspaper publishing companies, and local bulletins and newsletters. The local mass media was found to be quite effective in advertising this program. Moreover, past research results (described in section 6.6 in this chapter) on multiples were rewritten with advice for mothers of multiples so the results could be easily understood, and these results were provided in a fact sheet, brochure and leaflet presented to participants in workshops, family support events and similar settings. Finally, the '*Ishikawa Support Network for Multiple Birth Families*' was founded in July, 2005.

The current work was the first attempt in Japan to construct a support network system focused on the multiple birth families at the prefectural level. This network comprises a wide range of core members, including leaders of the twin mothers' clubs, maternal and child health professionals of governmental and medical institutions as well as from the private sector, and researchers. Its aims are to hold workshops, family support events and parenting classes specialized for families with multiples, to facilitate the exchange of information, skills and discussion on maternal and child health policies, and to promote total research on multiple births. The leaflet produced by this network is available in every PHC, municipal PHC and obstetric and pediatric institution, as well as in other places where it would come to the attention of expectant mothers or parents of multiples. It is important to introduce the network to the families of multiples. Public health nurses, who introduce the network during their home visits, accomplish this effectively, especially in rural areas. In some cases, mothers with experience in parenting multiples also visit maternity hospitals.

7.5 Peer support program in '*Ishikawa Support Network for Multiple Birth Families*'

A peer support program was started in 2007. This is a home-visit program involving well-trained peer supporters, namely the mothers of multiples. Its main differences with the previous family support provided by local twin mothers' clubs are as follows. First, the roles of coordinator (well-trained mothers) and peer supporter were clearly divided. Coordinators arrange requests from clients (mothers or expectant mothers of multiples), make support plans, and allot proper peer supporters according to the chief complaints of clients. They usually go along with peer supporter while home-visit, sometimes connect clients to social resources or health professionals of governmental and medical institutions. They also follow up the client and provide mental support for the peer supporter. The main role of the peer supporter is attentive listening to the client, and she also makes an activity log for case conferences. The peer supporter may advise the client, but does not impress her own opinion. Second, an education program by professionals is provided for the coordinator and peer supporter. This education program involves training in basic knowledge on multiple pregnancy and births, parenting multiples, and social resources and training in attentive listening. Third, since this program is considered one of the activities of the support network, it also incorporates training in systems dealing with clients who are having difficulty or who have emergency needs. Periodical meetings for the discussion of anecdotal reports or case conferences are held. Although it is very difficult to improve the statistical indicators of maternal and child health quickly, the questionnaire survey showed that the empowerment effect was observed not only for clients but also for peer supporters and coordinators. Moreover, the network itself was empowered by this activity.

7.6 Multiplied effect of this support network

By constructing a support network at the prefectural level, continuous support, an expansion of cooperation, a wide range of information exchange, support that is carefully crafted according to the characteristics of the region, a reduction of the overburden for specific key persons and expansible support became possible. By belonging to this network, local twin mothers' clubs gain avenues of communication with governmental or medical institutions. Moreover, the methodologies of social support networks in other fields with more members and longer histories, for example, patients' associations, are applicable.

Other prefectures have constructed or have been constructing and appear to be considering this kind of support network. The Ishikawa support network provides a good model. Although the method varies according to the real situation in a given prefecture, such as human resources or collaboration with governmental/medical institutions, the idea of a population approach is constant. The experience of the success or failure of the process of constructing a support network will be useful directly or indirectly to other prefectures. This strategy is welcomed by the many participants and other involved parties. Finally, the Japan Multiple Births Association (JAMBA) was established on February 22, 2010 (2/22/Heisei 22 in a Japanese era), in celebration of families with multiples. This association became a member of ICOMBO on June 2010.

8. Conclusion

Previous support systems were mainly focused on the individual or on small numbers of families with urgent needs or those willing to participate in child care classes. This is an approach to resolving problems downstream. However, many problems surrounding multiple birth families are wider social concerns. These problems have more essential background factors such as societal perspectives on fertility treatment, medical economics, perinatal medical systems, the institution of laws and guidelines, and human and social resources.

It is necessary to take an approach that is focused on the background factors and wider context to resolve these problems. Most data on multiples after birth can only be obtained by volunteer families with multiples. Epidemiologic studies should be performed based on community networks of multiple families and should reflect their real needs and unknown social concerns with appropriate and useful feedback based on scientific evidence, as shown in Figure 9. If research which actively involves the multiple birth families concerned and other community members is performed with CBPR approach, both families and researchers can benefit, and good relationships will be developed. The concept of the community support network seems to be one effective means of providing support for the multiple birth families. Evidence-based support for multiple birth families (EBPH in a broad sense) should be performed. Public health initiatives to resolve the many problems related to the rapid increase of multiple births are expected to be proposed and implemented.

In conclusion, the present approach of connecting research and practice can become an effective method applicable to other public health problems related to health inequality.

9. Acknowledgments

The authors gratefully acknowledge the help of research assistant Toshimi Ooma in the analysis of the data. I also thank to all board members of Ishikawa Support Network for Multiple Birth Families and Japan Multiple Births Association. This work was supported in part by a Grant-in-Aid for Scientific Research (B) from Japan Society for the Promotion of Science (Grant Number 21390206).

10. References

Akerman, B. A. & Fischbein, S. (1992). Within-pair similarity in MZ and DZ twins from birth to eighteen years of age. *Acta Genet Med Gemellol*, Vol.41, No.2-3, pp.155-164, ISSN 0001-5660

Blondel, B., Kogan, M. D., Alexander, G. R., Dattani, N., Kramer, M. S., Macfarlane, A. & Wen, S. W. (2002). The impact of the increasing number of multiple births on the rates of preterm birth and low birthweight: an international study. *Am J Public Health*, Vol.92, No.8, (August 2002), pp.1323-1330, ISSN 0090-0036

Blondel, B., Macfarlane, A., Gissler, M., Breart, G., Zeitlin, J. & PERISTAT Study Group. (2006). Preterm birth and multiple pregnancy in European countries participating in the PERISTAT project. *BJOG*, Vol.113, No.5, (May 2006), pp.528-535, ISSN 1470-0328

Bryan, E. (2003). The impact of multiple preterm births on the family. *BJOG*, Vol.110, Suppl.20, (April 2003), pp.24-28, ISSN 1470-0328

Bryan, E. (2006). Best Practice Guidelines. *Early Hum Dev*, Vol.82, No.6, (June 2006), pp.353-354, ISSN 0378-3782

Bryan, E., Denton, J. & Hallett, F. (1997). *Guidelines for Professionals: Multiple Pregnancy*, Multiple Births Foundation, London

Chaudhari, S., Bhalerao, M. R., Vaidya, U., Pandit, A. & Nene, U. (1997). Growth and development of twins compared with singletons at ages one and four years. *Indian Pediatr*, Vol.34, No.12, (December 1997), pp.1081-1086, ISSN 0019-6061

Denton, J. (2005a). Twins and more--1. Some current thinking on multiple births. *J Fam Health Care*, Vol.15, No.5, pp.143-146, ISSN 1474-9114

Denton, J. (2005b). Twins and more--2. Practical aspects of parenting in the early years. *J Fam Health Care*, Vol.15, No.6, pp.173-176, ISSN 1474-9114

Derom, C., Derom, R., Vlietinck, R., Maes, H. & Van den Berghe, H. (1993). Iatrogenic multiple pregnancies in East Flanders, Belgium. *Fertil Steril*, Vol.60, No.3, (September 1993), pp.493-496, ISSN 0015-0282

Derom, R., Bryan, E., Derom, C., Keith, L. & Vlietinck, R. (2001). Twins, chorionicity and zygosity. *Twin Res*, Vol.4, No.3, (June 2001), pp.134-136, ISSN 1369-0523

Flidel-Rimon, O. & Shinwell, E. S. (2006). Breast feeding twins and high multiples. *Arch Dis Child Fetal Neonatal Ed*, Vol.91, No.5, (September 2006), pp.F377-380, ISSN 1359-2998

Glinianaia, S. V., Skjaerven, R. & Magnus, P. (2000). Birthweight percentiles by gestational age in multiple births. A population-based study of Norwegian twins and triplets. *Acta Obstet Gynecol Scand*, Vol.79, No.6, (June 2000), pp.450-458, ISSN 0001-6349

Gray, M. (2008). *Evidence-Based Health Care and Public Health: How to Make Decisions About Health Services and Public Healt* (3nd), Churchill Livingstone, ISBN 978-0-44310-123-6, Edinburgh, UK

Gromada, K. K. & Spangler, A. K. (1998). Breastfeeding twins and higher-order multiples. *J Obstet Gynecol Neonatal Nurs*, Vol.27, No.4, (July-August 1998), pp.441-449, ISSN 0884-2175

Hall, J. E. & Callahan, T. L. (2005). 107 Economic Considerations. In: *Multiple Pregnancy: Epidemiology, Gestation & Perinatal Outcome* (2nd), Blickstein, I. & Keith, L. G. (Eds.), pp.889-894, Taylor & Francis, ISBN 978-1-84214-239-4, New York

Israel, B. A., Schulz, A. J., Parker, E. A., Becker, A. B., Allen, A. J., 3rd. & Guzman, R. (2008). Critical issues in developing and following CBPR principles. In: *Community-Based Participatory Research for Health: From Process to Outcomes*, Minkler, M. & Wallerstein, N. (Eds.), pp.47-66, Jossey-Bass, ISBN 978-0-47026-043-2, San Francisco

Jena, A. B., Goldman, D. P. & Joyce, G. (2011). Association between the birth of twins and parental divorce. *Obstet Gynecol*, Vol.117, No.4, (April 2011), pp.892-897, ISSN 0029-7844

Kato, N. (2004). Reference birthweight range for multiple birth neonates in Japan. *BMC Pregnancy Childbirth*, Vol.4, No.1, (February 2004), pp.2, ISSN 1471-2393

Leonard, L. G. (2003). Breastfeeding rights of multiple birth families and guidelines for health professionals. *Twin Res*, Vol.6, No.1, (February 2003), pp.34-45, ISSN 1369-0523

Leonard, L. G. & Denton, J. (2006). Preparation for parenting multiple birth children. *Early Hum Dev*, Vol.82, No.6, (June 2006), pp.371-378, ISSN 0378-3782

Machin, G. A. (1994). Twins and their zygosity. *Lancet*, Vol.343, No.8912, (June 1994), pp.1577, ISSN 0140-6736

Ooki, S. (2005). Language development of Japanese twins in childhood based on maternal reports. *Jap J Health and Human Ecology*, Vol.71, No.1, pp.12-24 ISSN 0368-9395

Ooki, S. (2006a). Motor development of Japanese twins in childhood as reported by mothers. *Environ Health Prev Med*, Vol.11, No.2, (March 2006), pp.55-64, ISSN 1342-078X

Ooki, S. (2006b). Population-based database of multiples in childhood of Ishikawa prefecture, Japan. *Twin Res Hum Genet*, Vol.9, No.6, (December 2006), pp.832-837, ISSN 1832-4274

Ooki, S. (2008). Breast-feeding rates and related maternal and infants' obstetric factors in Japanese twins. *Environ Health Prev Med*, Vol.13, No.4, (July 2008), pp.187-197, ISSN 1342-078X

Ooki, S. (2009). Chapter12. Strategy and practice of support for multiple birth families: Evidence-based care and population approach with human network, In: *Handbook of Parenting: Styles, Stresses, and Strategies*, Pacey, H., Krause, Tahlia, M., Dailey. (Eds.), pp.175-239, Nova Science Publishers, ISBN 978-1-60741-766-8, New York

Ooki, S. (2010). The effect of an increase in the rate of multiple births on low-birth-weight and preterm deliveries during 1975-2008. *J Epidemiol*, Vol.20, No.6, (November 2010), pp.480-488, ISSN 0917-5040

Ooki, S. (2011a). Effect of maternal age and fertility treatment on the increase in multiple births in Japan: vital statistics, 1974-2009. *J Epidemiol*, Vol.21, No.6, (November 2011), pp.507-511, ISSN 0917-5040

Ooki, S. (2011b). Estimation of the contribution of assisted and non-assisted reproductive technology fertility treatments to multiple births during the past 30 years in Japan: 1979-2008. *Twin Res Hum Genet*, Vol.14, No.5, (October 2011), pp.476-483, ISSN 1832-4274

Ooki, S. (2011c). Birth defects in singleton versus multiple ART births in Japan (2004-2008). *J Pregnancy,* Vol.2012, 2012:285706, ISSN 2090-2727

Ooki, S. & Asaka, A. (1993). Physical growth of Japanese twins. *Acta Genet Med Gemellol*, Vol.42, No.3-4, pp.275-287, ISSN 0001-5660

Ooki, S. & Asaka, A. (2004). Zygosity diagnosis in young twins by questionnaire for twins' mothers and twins' self-reports, *Twin Res*, Vol.7, No.1, (February 2004), pp.5-12, ISSN 1369-0523

Ooki, S. & Asaka, A. (2005). Comparison of obstetric and birthweight characteristics between the two largest databases of Japanese twins measured in childhood. *Twin Res Hum Genet*, Vol.8, No.1, (February 2005), pp.63-68, ISSN 1832-4274

Ooki, S. & Asaka, A. (2006). Twin database of the secondary school attached to the faculty of education of the university of Tokyo. *Twin Res Hum Genet*, Vol.9, No.6, (December 2006), pp.827-831, ISSN 1832-4274

Ooki, S., Yamada, K. & Asaka, A. (1993). Zygosity diagnosis of twins by questionnaire for twins' mothers. *Acta Genet Med Gemellol*, Vol.42, No.1, pp.17-22, ISSN 0001-5660

Ooki, S. & Yokoyama, Y. (2003). Reference birth weight, length, chest circumference, and head circumference by gestational age in Japanese twins. *J Epidemiol*, Vol.13, No.6, (November 2003), pp.333-341, ISSN 0917-5040

Ooki, S. & Yokoyama, Y. (2004). Physical growth charts from birth to six years of age in Japanese twins. *J Epidemiol*, Vol.14, No.5, (September 2004), pp.151-160, ISSN 0917-5040

Ooki, S., Yokoyama, Y. & Asaka, A. (2004). Zygosity misclassification of twins at birth in Japan. *Twin Res*, Vol.7, No.3, (June 2004), pp.228-232, ISSN 1369-0523

Peter, I., Vainder, M. & Livshits, G. (1999). Genetic analysis of motor milestones attainment in early childhood. *Twin Res*, Vol.2, No.1, (March 1999), pp.1-9, ISSN 1369-0523

Rietveld, M. J., van Der, Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E. & Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Res*, Vol.3, No.3, (September 2000), pp.134-141, ISSN 1369-0523

Rose, G. (1994). *The Strategy of Preventive Medicine*, Oxford University Press, ISBN 978-0-19262-486-4, New York

Rutter, M., Thorpe, K., Greenwood, R., Northstone, K. & Golding, J. (2003). Twins as a natural experiment to study the causes of mild language delay: I: Design; twin-singleton differences in language, and obstetric risks. *J Child Psychol Psychiatry*, Vol.44, No.3, (March 2003), pp.326-341, ISSN 0021-9630

Tanimura, M., Matsui, I. & Kobayashi, N. (1990). Child abuse of one of a pair of twins in Japan. *Lancet*, Vol.336, No.8726, (November 1990), pp.1298-1299, ISSN 0140-6736

Thorpe, K., Golding, J., MacGillivray, I. & Greenwood, R. (1991). Comparison of prevalence of depression in mothers of twins and mothers of singletons. *BMJ*, Vol.302, No.6781, (April 1991), pp.875-878, ISSN 0959-8138

Thorpe, K., Rutter, M. & Greenwood, R. (2003). Twins as a natural experiment to study the causes of mild language delay: II: Family interaction risk factors. *J Child Psychol Psychiatry*, Vol.44, No.3, (March 2003), pp.342-355, ISSN 0021-9630

Wilson, R. S. (1986). Growth and development of human twins. In: *Human Growth* (Vol.3 2nd), Falkner, F. & Tanner, J. M. (Eds.), pp.197-211, Plenum Press, New York

Section 4

Special Populations and Settings

20

TB Control in Prisons

J. Noeske
Health/AIDS Programme Cameroon,
Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ)
Cameroon

1. Introduction

According to WHO, in 2010, there were 8.8 million (range, 8.5-9.2 million) incident cases of TB and 1.1 million (range, 8.5-9.2 million) deaths from TB among HIV-negative cases and an additional 0.35 million (range, 0.32-0.39) deaths from HIV-associated TB (WHO 2011). An examination of the map of estimated incidence rates by country and the breakdown of national figures in social strata rapidly reveals that global statistics conceal profound inequalities in the present intra- and inter-country distribution of the disease. Ninety-five percent of the global TB disease and TB death burden is found in poor countries. In highly industrialized countries, TB is currently limited to socially marginalized and other poor high-risk groups such as IV drug users, migrants from developing countries, and, over the past 30 years, HIV-infected persons.

Prison inmates[1] constitute a high risk-group for tuberculosis (TB) in both developing and the industrialized countries. Infection and disease rates remain much higher – from 5 up to more than 80 times – compared to national averages (Conninx R et al. 1995; Aerts et al. 2006; Dara et al. 2009). In their majority, inmates originate from marginalized populations characterized by poor socio-economic living conditions and associated poor life-styles, the over-representation *per se* of multiple high-risk groups already being a significant factor in explaining the markedly higher TB incidence and prevalence in prison populations. Greater population susceptibility include risk factors as urban dwelling, crowding, poor housing with lack of ventilation, limited access to health services and, at individual level, smoking and alcohol use, HIV-infection, exposure to indoor air pollution, diabetes mellitus, and under-nutrition (Rieder 1999; Lönnroth et al. 2009, 2010; Murray et al. 2011). Imprisonment conditions, globally worsening along a downward gradient of per capita gross domestic product, amplify those risk factors. Prisons tend to be overcrowded with poorly ventilated cells. The proximity of large numbers of individuals for an extended time in such conditions facilitates TB transmission. Moreover, prison populations often suffer from mal-nutrition or under-nutrition, in particular in low- and low-middle income countries. Also, the latter sometimes bear the burden of high or intermediate TB endemicity and generalized HIV-endemics with penitentiary health services being inexistent, sub-standard or difficult to access. Prisons thus constitute an ideal environment for TB transmission.

---

[1] The term « prison » refers in this text to all places of detention, jails (pre-trial detention) as well as prisons (detention of sentenced prisoners). The terms "inmates" or "prisoners" are used alternatively, referring to all categories of prisoners, sentenced prisoners, remand prisoners, female, male and juvenile prisoners.

TB in prisons does not only affect prisoners (but also prison guards or visitors). An habitually important yearly turn-over of prison populations make prisons a reservoir for TB transmission to the community at large. TB is spread by released prisoners who had been infected by TB in prison and are at risk of progression from latent TB infection (LTBI) to TB disease, prisoners with TB disease not diagnosed in prison, those under treatment but still not cured, and prisoners with sub-standard treatment. The way TB in prison influences TB and TB control in general populations has been explored by recent research. Stuckler et al. (2008) established a clear relationship between rises in incarceration rates and increased TB incidence and multi-drug resistant TB prevalence rates in Eastern European and Central Asian countries. Increases in the size of prison populations accounted for a 20.5% increase in TB incidence in the countries studied from 1991 to 2002. Based on these observations and on the fact that patterns of communicable diseases are thought to be influenced by the existence of pockets of high transmission, Basu et al. (2011) modeled the role of what they referred to as "institutional amplifiers" in the dynamics and control of tuberculosis epidemics. Prisons play the role of an "institutional amplifier" of TB with, similarly like, mines and communal hospital wards among others. Using this model, they demonstrated that even a substantial increase in case detection and treatment success rates, the traditional control measures, has little effect on the overall population incidence as long as hubs of transmission continue to exist. According to these authors, a key intervention to reduce TB incidence, prevalence and mortality in a community would consist in limiting the number of persons entering institutional amplifiers.

This preliminary remark shows that neglecting TB prevention and control in prison settings could have serious health consequences for both prisoners and the civilian population, in particular in those in middle-low and low income countries where incarceration rates are high, where TB control is poor, and where HIV infection rates are high.

2. The burden of TB in prisons

Many observational studies in the USA and Eastern Europe have reported the burden of TB in prisons, although they are largely heterogeneous due to differences in study design, study populations, sampling methods, and the quality of their methodology. Retrospective and/or point prevalence studies prevail. They document prevalence of TB infection and disease in new prison entrants, exposure to and transmission of TB infection and progression to TB disease within prison populations including prison guards, as well as transmission of TB from within prisons to extramural civilian populations. In the late 1970s, the discovery of two cases of infectious TB in a State prison in Arkansas in the US prompted a careful study evidencing rapid intramural spread of the infection, higher morbidity due to TB in prisons in compared to the general population and the transmission of the disease from a former inmate to individuals outside the prison (Stead 1978). With the transient rising incidence of TB in the late 1980s in the US, caused by prematurely dismantled TB services and fuelled, in parallel, by the emerging HIV epidemic, TB among prison populations received increased attention by researchers. The yearly number of published studies concerning TB in prisons as reported by MedLine and other databases started increasing. In 1989, a dramatic increase in the incidence of TB in prisons in New York State was reported with over half of the prisoners with the disease also having AIDS or infected with HIV. (Braun et al. 1989; Darbyshire 1989) Outbreaks of multi-drug-resistant TB in New York State prisons in 1990/91 alerted US public health authorities and the international community (Valway et al. 1993, 1994; Drobniewski 1995). Nosocomial (MDR-)TB transmission from inmates to prison staff was reported from a California prison by CDC in 1993 (CDC 1993). An initial survey of TB incidence in a large number of States in nursing homes and correctional facilities in the US showed the aggregate TB incidence rate for inmates in correctional facilities to be 3.9 (95%CI: 3.35-4.49) times higher than the rate for persons of a similar age not incarcerated, thus more systematically confirming former observational evidence from individual prisons. (Hutton 1993) Other reports from the US which followed confirmed these findings. Increased public funding for TB control efforts and the consequent implementation of direct observed therapy (DOT) in high incidence areas and populations reversed TB incidence rates from 1993 onwards. In jails and prisons, however, better controlled TB also remained a major public health concern (Kendig 1998; MacNeil et al. 2005). Subsequent studies concentrated on specific determinants and risk factors such as increasing TB infection rates associated with increased time spent in the jail system (Bellin et al. 1993), HIV-infection and TB among risk groups such as intravenous drug users (IDU) (Pelletier et al. 1993), molecular demonstration on intramural transmission (Valway 1991; Ijaz et al. 2004), or the evaluation of control measures (Curtis et al. 1994; MacIntyre et al. 1997; Klopf 1998; CDC 2006). Reports on TB in prisons of Western European countries experiencing an evolution similar to that of the US soon confirmed the US findings. Studies conducted in Spanish prisons reported high prevalence of TB infection and illness upon admission in jail, described risk factors as HIV-infection, IDU status, malnutrition, or immigrant status. TB transmission in prisons, first described by conventional epidemiological methods, was underpinned by molecular techniques from of the 1990s. March et al. (2000) described 14 unsuspected active chains of TB transmission involving 65 patients in Barcelona prisons system combining conventional epidemiological techniques with DNA fingerprinting of *M. tuberculosis*. Recent transmission accounted for almost half of the incident cases, confirming observations made in Madrid prison population during the same period (Martin et al. 1994; Chaves F et al. 1997; Fernandez-Martin et al. 2000). In 2000, a study evaluated high TB incidence rates using molecular techniques in French prisons (Hanau-Berçot et al. 2000). In 2002, a systematic data collection by questionnaire in 52 European countries with 22 countries responding revealed a median TB notification rate of 232 per 100 000 inmates (0-17 808), highlighting the vulnerability of prisoners to TB, with the prison populations of Eastern European countries being the most affected (Aerts et al. 2006).

After the disintegration of the former USSR with subsequent economic liberalization the downward trends in notification of tuberculosis observed from the 1950s to the 1990s reversed throughout all former Soviet Union States. At the same time, the average age of TB patients declined, reflecting high levels of recent transmission. From the 1990s onward, a growing number of studies reported the high burdens of TB, in particular MDR-TB, in prisoners in the former USSR Republics (Drobniewski 1995 ; Drobniewski et al. 1996; Conninx et al. 1998; Kimerling et al. 1999; Aerts et al. 2000). Although there is in recent year a tendency towards a reduction in TB incidence among prison inmates, a higher proportion of TB patients among prisoners have MDR strains than in patients outside prison. A challenge is, too, the rising number of people infected with HIV.

TB is a major public health problem in Africa which also has a high prevalence of HIV infection. Active case finding surveys from Ivory Coast and Malawi reported, for the first time, high prevalence and incidence rates of pulmonary TB in an African prison in the 1990s (Nyangulu et al. 1997; Koffi et al. 1997). Cross-sectional studies from Tanzania (Rutta et al. 2001), Botswana (CDC Report 2003), Cameroun (Noeske et al. 2006) and others followed, all confirming high TB prevalence rates among inmates and suggesting TB prevalence being several fold higher in prison than among the civilian population. A review of all relevant English publications on TB in prisons in Sub-Saharan Africa (SSA) performed by O'Grady et al. (2011) concluded that there is evidence of an increasing prevalence of active TB in prisons in SSA with drug resistant TB increasingly being detected. In terms of the driving factors, the study identified HIV, living conditions, including overcrowding and poorly ventilated cells, as well as malnutrition associated with low immunity. Concerning the root causes of this situation, the authors identified lack of public funding and mismanagement of meager resources. In addition, prison health services were found to be insufficient.

Reports and studies from other continents confirmed the overall trend. Ferreira et al. reported high incidence rates of TB among incarcerated female prisoners, HIV-infection being a major risk factor in a Brazilian prison (1998). The observations were confirmed for other prison populations by study results from a Rio de Janeiro prison (Sanchez et al. 2005). A TB survey in Taiwan during 1998/99 based on mass radiographic screening showed high TB prevalence in all 29 jails and prisons studied (Chiang et al. 2002). TB prevalence and incidence among Thai prison inmates were shown to exceed by large those in the comparable civilian population (Sretrirutchai et al. 2002; Jittimanee et al. 2007). In the Karachi central prison, TB appeared to be 3.75 times higher than in the general population of the town (Rao 2004). A TB surveillance study in correctional institutions in Hong Kong over the period 1999-2005 found very high TB prevalence among prisoners before and within the first months of incarceration (Wong et al. 2008).

In spite the overwhelming published evidence of TB in prisons as a permanent and, perhaps, even a growing problem, some significant gaps persist on our understanding of TB in prisons as highlighted in the editorial of a recently published systematic review on studies evaluating the incidence of latent tuberculosis infection and TB in prisons worldwide and among the general population attributable to the risk of prison-to-community TB propagation (Baussano et al. 2010). It is still unclear how much TB there is in prisons, and it is not possible to give accurate estimates of the increased incidence of TB over community rates; the impact of specific prison conditions on TB transmission is yet to be evaluated; as is the fraction of TB among the general population due to disease among incarcerated groups.

3. Epidemiology of TB in prisons

Following the tuberculosis classification of the American Thoracic Society and the United States Centre for Disease Control (CDC, now Centre for Disease Control and Prevention), Rieder (1999) proposed a model allowing the distinction of four major stages for comprehending the dynamics of TB.

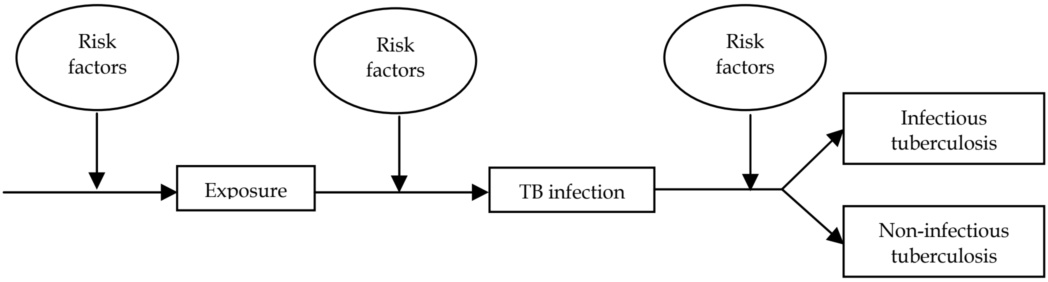


Fig. 1. A simplified model for tuberculosis epidemiology.
Figure adapted from Rieder (1999), p9.

Adapting this model to the conceptualization of the epidemiological framework for understanding the dynamics of tuberculosis in prisons, we will briefly discuss exposure to TB, infection with TB, and TB disease in prison before discussing TB control policies and strategies.

3.1 Exposure to tubercle bacilli in prison

The main factors determining the risks of exposure to tubercle bacilli are"… the number of incident infectious cases, the duration of their infectiousness and the number and nature of interactions between a case and a susceptible contact per unit of time of infectiousness" (Rieder (1999), p11). Concerning exposure to tubercle bacilli in prisons, the following practical observations could be made: (1) Empirically, a significant number of infectious cases is generally present in prison settings; (2) such cases tend to be infectious for a prolonged duration, either because they are not diagnosed (at entrance, during stay) or they are diagnosed late or are not treated at all treated or are poorly-treated; (3) infectious cases are in close and prolonged human contact in overcrowded and poorly ventilated cells.

3.2 Infection with tubercle bacilli in prison

The infectious bacterial disease is caused by M. tuberculosis. Humans are the principal reservoir for the *M. tuberculosis* complex with the pathogenic species *M.tuberculosis*, *M. bovis*, and *M. africanum,* the first being the most frequent cause of TB disease. Most commonly, TB affects the lungs. TB is an almost exclusively airborne infection caused by tubercle bacilli which spreads as infectious droplet nuclei. The link between the transmission of TB infection and overcrowding is a well-known phenomenon since the exploration of aerobiology in the 1950s and 1960s by Wells, Riley and their co-workers. Beggs et al. (2003) reviewed alternative epidemiological models representing the transmission of tuberculosis in confined spaces. The authors identified the following factors as crucial for effective TB transmission in confined spaces: long exposure periods, sufficient number of infectious droplet, small room volume, high occupancy density, low ventilation rate, proximity to the infector, and the infectiousness of the incident case. However, the number of infectious nuclei required to infect humans with any airborne infection is unknown. Regarding the infection of humans to become infected, it is difficult to measures factors like the virulence of the TB strain and the resistance of infected individual comes into play, as well as factors involved in the dynamics of the infection.

There are currently two main tools available for determining infection with MTB, namely: the tuberculin skin test (TST) and interferon-gamma release assays (IGRAs). Tuberculin, a concoction of antigens produced from killed tubercle bacilli, was first introduced by Koch. Subsequently, to increase the specificity of the tuberculin test was purificated, and the dose and administration technique standardized. Depending on the epidemiological situation (TB prevalence, presence of atypical mycobacteria, presence of HIV infection or other immune-compromising factors), the diameter of TST induration can indicate the probability of TB infection. Over the past decade two commercial have been developed and FDA-approved. IGRAs have become widely endorsed in high-income countries for diagnosis of latent TB infection (LTBI) and several guidelines on their use have been issued (Masurek et al. 2010). However, following a very recent Policy Statement, WHO concludes that there is insufficient data and low quality evidence on the performance of IGRAs in low- and middle-low income countries, typically those with a high TB and/or HIV burden, that IGRAs are more costly and technically complex to conduct than the TST and, as such, replacing TST by IGRAs as a public health intervention in resource-constrained settings is not recommended (WHO 2011). An overall conclusion can be that neither IGRAs nor TST can accurately predict the risk of infected individuals developing active TB disease nor that they can be used for the diagnosis of active TB; that in prison settings in high-income countries, IGRAs can play a key role in the diagnosis of latent TB infection; that TST in prisons in low and middle-low income countries is of limited value given the overall TB infection rate of the population and the HIV infection burden. Studies suggest that HIV-infected individuals are more likely to be infected after exposure to MTB (Daley et al. 1992).

In their systematic review of available published evidence on incident latent tuberculosis infection (LTBI) Baussano et al. (2010) identified six studies, 5 from the US and one from Brazil, all dating more than ten years. The median estimated annual incidence rate ratio (IRR) for LTBI was 26.4 (interquartile range: 13.0-61.8). The authors also estimated the fraction of LTBI in the general population attributable (PAF) to within-prison exposure to MTB for each study. The median estimated PAF for LTBI was 13.1% (8.1%-30.0%) for high-income countries and 6.3% (2.7%-17.2%) for studies from middle-low-income countries. These findings suggest that the within-prison spread of LTBI is likely to substantially affect the incidence of LTBI and TB in the general population.

3.3 Progression to tuberculosis disease in prisons

An estimated one-third of the human population is infected with MTB. Most of it does not develop TB disease. Why a particular person does or does not develop TB disease after having been infected by the tubercle bacilli is difficult to determine. On the other hand, a series of factors have been identified as being are more or less strongly associated with progression to TB disease once infected with MTB. Rieder (1999, p64) summarizes a selection of these risk factors together with their strength of association with the disease. Prison inmates are in particular concerned by the following among those factors (according to the strength of association): HIV infection, TB infection acquired less than one year ago, underweight and the quantity of the infecting dose. The first two factors have been identified in population based studies and the other two observed in numerous case-control studies. The strongest risk factor for developing TB is infection with HIV. The risk of disease after infection is 10 percent per year among people living with HIV without ART, compared to 10 percent per lifetime among those who are negative. There is evidence that TB infection among HIV-infected patients progresses to TB more rapidly than among those without HIV infection. Habitually, inmates are characterized not only by one, but several factors at a time.

How can the magnitude of the TB problem in prison populations be quantified? Rieder et al. (2011) discuss in a recent paper the critically different calculations for epidemiological indicators of the TB burden in prisons. Case definition, depending on screening criteria and length of stay in prison, are as critical as the use of appropriate numerator and denominator data. For measuring TB prevalence the authors propose medical screening including TB screening at entry with clear case definition and appropriate bacteriological examinations, discarding cases diagnosed in the first 3 months of stay as cases prevalent at entry. Using a simple fictitious data set, the authors show the limitation of cross-sectional prevalence surveys as an instrument for assessing time trends in prevalence. For measuring TB incidence rate (person-time of observation at risk of developing the disease), the most appropriate measure for estimating the magnitude of the TB problem in a prison setting, the authors propose as denominator the exact date of entry, excluding the first 3 months, and the exact date of exit (occurrence of TB, release, end of the observation period).

Baussano et al. in their above-cited study identified only nineteen available published studies answering the criteria for assessing the risk of incident TB disease within prisons. They calculated the estimated median annual IRR for TB which was 23.0 (11.7-36.1). However, all but three of the studies retained for the review were dated from during the last century and only two had been carried out in low(-middle) income countries (Ivory Coast, Brazil), confirming the need to sustain research efforts in view of further up-to-date and accurate estimates on the increased incidence of TB in prisons.

4. TB control in prisons[2]

The control of TB in prisons and other congregate settings has to take in account the entrance point prevalence of infection among institutional residents (and staff), the potential for reactivation, the role of transmission within the institution, the potential for detection of infection and disease, the potential for prevention and treatment of disease, and the potential of the building environment to favor transmission (Nardell 1989). Effective chemotherapy is the single most important factor in reducing infectiousness in any situation and at any stage of the disease. This requires an effective TB control program with early case finding and subsequent successful treatment. Since the early 1990s, the principles of prevention, diagnosis, treatment, and care of TB have been codified by WHO in the so-called DOTS strategy (DOTS=directly observed treatment, short course). Key elements are political commitment with increased und sustained financing; case-detection through quality assured bacteriology; standardized treatment with supervision and patient support; an effective pharmaceutical supply and management system; and a monitoring and evaluation system and impact measurement. The DOTS strategy was expanded in 2006 to The Stop TB Strategy (WHO 2010a) and launched by the Stop Tuberculosis Partnership in a Global Plan to Stop TB in 2010 (WHO 2010b). Its objectives and components apply, in

---

[2] Cf. for the following: WHO 2007a

general, to TB control in prisons as well. And corresponding comprehensive guidelines on how to control and prevent TB in prisons are available since 1998 (Maher et al. 1998; Bone et al. 2000; Bock 2000; WHO 2007a; Dara et al. 2009). These guidelines, however, are seldom implemented in developing countries. And even in developed countries the strict and competent application of this recommendation is still hampered. Sosa et al. (2006) reported a tuberculosis outbreak in a correctional facility after two recently discharged inmates were diagnosed with TB. Despite prolonged symptoms and abnormal chest radiographs, the index case was not diagnosed while incarcerated and fifty three presumably exposed inmates and ten of 485 correctional staff had a conversion of their TST. – Taking as guiding principles the elements of the DOTS strategy and taking in account the six components of The Stop TB Strategy the following text will resume essentials of TB control strategies in prisons. The above cited guidelines provide the necessary additional technical details.

4.1 Political commitment with increased and sustained financing

Health care services and national TB services commonly are under the responsibility of a ministry of (public) health. Prison health care mostly is under a different ministry (Justice, Internal Affairs or other). For these latter ministries in charge, in particular in low- and middle-income countries, prison health stands seldom on a priority list. Consequently, prison health services are as a rule grossly underfunded. This underfunding influences negatively all components of the prison health system, the service delivery, the health workforce, infrastructures and medical products and technologies, the information system, financing of the users, the technical leadership and the overall governance. TB control cannot act as the main leverage for changing fundamentally this situation. National TB programmes (NTP), however, are generally well-structured and dispose of detailed technical guidelines. Preventive, diagnostic and treatment procedures and means, supply of drugs and laboratory consumables, recording and reporting tools, indicators for monitoring and evaluation are well defined. In addition, the principle of equivalence requires that prisoners have access to the standards of health care as the general population. Moreover, most PNTs in high-burden countries benefit today from external funding. Experience has clearly shown that prison TB programmes should be coordinated with and integrated in civil-sector TB programmes. Improved collaboration or integration contributes to strengthening the overall health system and make benefit a habitually weak prison health system. Linkage to the civilian sector has the additional advantage of facilitating treatment follow-up of release prisoners.

4.2 Case detection, including control of MDR and HIV

The DOTS strategy traditionally relies on passive case finding, meaning that TB cases are waited for presenting themselves spontaneously to health services in order to get diagnosed and treated. International guidelines (WHO 2007b) recommend on entry into prison a medical examination that should include TB screening. We saw that prison inmates belong rather to minorities or migrant groups or other segments of the population in which TB occurs with a higher incidence. The implementation of active case finding strategies should be standard. - One is the screening of new prison entries for active TB disease; this recommendation is valid or all type of prisons. Various screening tools have been proposed and assessed, varying from symptom based (clinical) screening including anamnesis with regard to past or present TB treatment, over chest radiography, TST or IGRA, or a combination of this methods. What screening tools are used has to depend on NTP guidelines, local feasibility, and available means. Results of clinical scores may depend on the setting were they are applied and should be tested and adapted. The diagnosis of pulmonary TB (PTB) cannot be based neither on the result of chest radiography nor of TST or IGRA. According to WHO recommendations, all PTB B suspects should undergo appropriate bacteriological sputum examination. Sputum negative PTB suspects should be diagnosed according to the national NTP guidelines. An incomplete network of smear microscopy sites inside the prison system can be completed by collaboration with the civil-sector TB microscopy network. - The usefulness of periodic screening during the stay in prison is much debated, but its implementation seems unavoidable in high TB incidence prisons or compartments of prisons. Mathematical modeling of the TB epidemiological dynamic according to control strategy shows that also passive TB detection may remain a priority in TB control in general populations, it appears insufficient in highly TB-endemic prisons to produce a rapid decline in TB incidence (Legrand et al. 2008). Likewise, Uys et al. (2011) calculated TB transmission possibilities in low and high prevalence areas and settings and came to the conclusion that in high prevalence settings transmission is relatively unresponsive to changes in the number of infectious people. He coins the term of "transmission elasticity" to describe this phenomenon and he warns against excessively optimistic projections regarding the effectiveness of 'habitual' control strategies. For TB control in prison settings he strongly recommends two measures: ventilation and consequent contact tracing. – Finally, contact investigation around a newly detected case within the prison has been recommended. The mathematical modeling is confirmed by an observation published by Noeske et al. (2011) according to which consequent TB screening at entrance and active contact tracing are not able to keep up with TB transmission dynamics in largely overcrowded and non-ventilated cell blocks.

In high-income low-incidence settings like in the US, TST or IGRA for documenting TB infection in inmates at entrance or during prison stay as well as in prison staff is recommended. In low-income high-burden settings, this testing for LTBI is of limited value as a majority of young individuals are TB-infected and/or have received BCG vaccination which makes difficult the interpretation of the test. However, where a national policy for TB prevention in HIV infected individuals exists, the test might have a place in a range of instruments applied for excluding active TB and deciding for isoniazid (INH) or similar TB prophylaxis.

Ideally, all TB suspects should in addition to sputum smear microscopy undergo drug susceptibility testing (DST). Again, implementation of this policy depends on the local feasibility and the available means. In any case, prison inmates have to have access to the same facilities as the general population.

Seen the close association between HIV and TB, provider initiated voluntary counselling and HIV testing should be an integral part of the prison's TB control programme. Ideally, to all prisoners who are screened for TB a voluntary HIV test should be proposed. For prisoners tested positive, the whole range of further diagnostic exploration and prophylactic and definite treatment (ART) should be made available and accessible.

4.3 Treatment with supervision and patient support

Treatment and management of TB obey the same principles in prison as in the civil sector. At first glance, directly observed treatment seems easily to maintain as prisoners are in one place and always present. However, when prison health personnel are lacking or not motivated, observation of adherence to treatment is difficult. Charging prisoners with dispensing TB drugs leaves the correct administration of medications to the *alea* of multiple informal hierarchies and power structures. Ensuring that every dose of drugs is taken directly observed remains a challenge.

Another challenge are prisoners released while still on treatment. Formal referral with forms is basic, but not sufficient. Good experiences have been made by combining sensitization of the patient for the necessity to complete treatment, an agreement together with the patient on where and how to complete treatment, and social workers – or NGO members – constituting a kind of mediator between patients, the prison health service and the civil TB dispensary network susceptible to receive the once released prisoner. This needs mostly external funding as do incentive packages which are used in some settings to motivate patients.

4.4 Information, education and communication (IEC) for prisoners and prison staff

The evidence base for information, education and communication (IEC) for prisoners and prison staff is relatively weak compared with that for the other elements of the DOTS strategy. Most studies conclude that the effectiveness of current educational efforts in influencing prisoners′ behavior remains largely unknown. In particular, studies have pointed out that education and counseling are not of much use to prisoners if they do not have the means to act on the information provided while they are in prison. Recognizing that prisons and other closed settings are important settings for IEC programs for both prisoners and staff about HIV and other infectious diseases, well-designed programs should be implemented in all prisons and other closed settings. Written materials should be appropriate for the educational level in the population in prisons. Furthermore, prisoners and staff should participate in the development of educational materials. Where possible, education delivered for prisoners by the prison system should be supplemented by peer education programs. In particular, prisoners must be provided with the prevention measures.

4.5 Drug supply and management system

The NTP must be made entirely responsible and capable of providing sufficient and adequate drugs and other consumables to cover the need of all patients in the civil sector as well as in prisons.

4.6 Monitoring, evaluation and impact measurement

The recording and reporting system in many low(-middle) income and high TB burden countries is weak. The causes for lack of motivation for prison health staff to report correctly are multiple: Lack of time because of personnel shortage, of essentials like registers, treatment cards, even stationary, together with poor training and supervision, and little use of data for analysis. Regularly, too, prison administration intervenes in the technical reporting process for political reasons (too many deaths shed a bad light on the administration). Where there TB registers, treatment cards and report instructions they might differ from those of the TB program. For epidemiological purposes and better follow-up of patients the recording and reporting system for TB in prison should be identical with that of the civil sector. When a prison is collaborating closely with a civil sector diagnostic and/or treatment center without self having his status, patient management tools should be doubled in order to allow the prison health staff to follow-up the prison's TB patients according to the national guidelines. Reporting should be done to the civil-sector TB-program and, if needed on the basis of a written agreement, supervision and evaluation should be done together with or in close coordination with the NTP.

4.7 Infection control

To reduce the risk of TB transmission by infectious patients to susceptible individuals (other prisoners, staff, and visitors), early diagnosis and prompt treatment, but also the separation of patients during the period of infectiousness is recommended. The consultation room should be well ventilated as should the wards reserved for infectious TB patients. Patients should be educated with regard to cough hygiene. Whether personal protection (respirators for staff, disposable masks for patients, staff, and visitors) and engineering measures (negative ventilation for example) can be used will depend on available financial means.

5. Conclusion

At present, about 10 million people are incarcerated world-wide. The majority among them is at increased risk for TB which is largely a curable disease. TB prison control guidelines recommend technical solutions mainly formulated in operational public health terms. However, even most comprehensive TB control programs cannot succeed unless there is a fundamental change in confinement conditions like overcrowding, poor ventilation, under-nutrition, but also informal hierarchies, corruption and violence with consequences for access to health services contribute decisively to maintain the vigorous and vicious circle of TB transmission in prisons. Until then, inmates are doubly punished, incarcerated and incessantly exposed to TB, their situation remaining a shame to public health and to human rights.

6. Annex

TB control programmes in prisons: What would be a minimum standard?

The ideal TB control programme in a prison would include the following:

- government structures supporting TB programmes in both the civil-sector and penitentiary systems;
- written agreement on collaboration and coordination between prison and civil-sector TB programmes;
- prevention through the early detection of infectious cases; avoidance of overcrowding; good nutrition, ventilation and light; clear definition of infectious zones with clear policies on how to reduce transmission; and the use of masks and/or personal respirators;

- complete access to TB diagnosis and treatment for all prisoners entering the prison system;
- adequate treatment in line with national TB programme guidelines, including those for prisoners with MDR-TB and TB-HIV coinfection, using DOTS and an uninterrupted supply of drugs of guaranteed quality; and
- a guarantee by prison and civil medical personnel of continued treatment for infected individuals following their release from prison.

The penitentiary system must have:

- continuing human resource development that ensures an adequate number of staff with satisfactory background education and continuous training;
- a mechanism in place for timely investigation of TB suspects and early detection of individuals with active TB and their treatment;
- a network of laboratories carrying our quality-assured smear microscopy;
- drug susceptibility testing in a centralized laboratory of ensured quality, either in the civil sector or in the prison;
- a supply of quality second-line drugs for prisoners suffering from MDR-TB once the capacity to test prisoners for drug resistance is in place; and
- recording and reporting in close coordination with the civil sector.

Source:

World Health Organization. Regional Office for Europe. Status Paper on Prisons and Tuberculosis. Copenhagen 2007 (p23).

7. References

Aerts A, Habouzit M, Mschiladze L, Malakmadze N, Sadradze N, Menteshashvili O, Portaels F, Sudre P. Pulmonary tuberculosis in prisons of the ex-USSR state Georgia: results of a nation-wide prevalence survey among sentenced inmates. Int J Tuberc Lung Dis. 2000; 4(12):1104-10.

Aerts A, Hauer B, Wanlin M, Veen J. Tuberculosis and tuberculosis control in European prisons. Int J Tuber Lung Dis. 2006; 10(11):1215-23.

Basu S, Stuckler D, McKee M. Addressing institutional amplifiers in the dynamics and control of tuberculosis epidemics. Am J Trop Med Hyg 2011; 84(1):30-7.

Baussano I, Williams BG, Nunn P, Beggiato M, Fedeli U et al. Tuberculosis incidence in prisons: A systematic review. 2010. PLoS Med 7: e 381.doi: 10.1371/journal;pmed.1000381 (accessed 15/12/2011).

Beggs CB, Noakes CJ, Sleigh PA, Fletcher LA, Siddiqi K. The transmission of tuberculosis in confined spaces: an analytical review of alternative epidemiological models. Int J Tuberc Lung Dis 2003; 7(11):1015-26.

Bock NN. Tuberculosis in Correctional Facilities. In: Reichmann LB, Hershfield ES (eds). Tuberculosis. A comprehensive International Approach. 2nd revised and expanded edition), New York, Basel: Marcel Dekker; 2000; p. 645-60.

Bone A, Aerts A, Grzemska M, Kimerling M, Kluge H, Levy M, Portaels F, Raviglioni M, Varaine F. Tuberculosis Control in Prisons. A Manual for Programme Managers.

Geneva 2000 (World Health Organization/ International Committee of the Red Cross: WHO/CDS/TB/2000.281).

Braun MM, Truman BI, Maguire B et al. Increasing incidence of tuberculosis in a prison inmate population : Associated with HIV infection. JAMA 1989; 261:393-397.

Centers for Disease Control and Prevention (CDC). Probable transmission of multidrug-resistant tuberculosis in a correctional facility--California. MMWR Morb Mortal Wkly Rep. 1993 Jan 29; 42(3):48-51.

Centers for Disease Control and Prevention (CDC). Rapid assessment of tuberculosis in a large prison system--Botswana, 2002. MMWR Morb Mortal Wkly Rep. 2003. Mar 28; 52(12):250-2.

CDC. Prevention and Control of Tuberculosis in Correctional and Detention Facilities. Recommendations from CDC. MMWR 2006; Vol.55 / No.RR-9.

Cegielski JR, McMurray DN. The relationship between malnutrition and tuberculosis: evidence from studies in humans and experimental animals. Int J Tuberc Lung Dis 2004; 8(4): 286-298.

Chaves F, Dronda F, Cave MD, Alonso-Sanz M, Gonzalez-Lopez A, Eisenach KD, Ortega A, lopez-Cubero L, Fernandez-Martin I, Catalan S, Bates JH. A longitudinal study of transmission of tuberculosis in a large prison population. Am J Respir Crit Care Med. 1997; 155(2):719-25.

Chiang CY, Hsu CJ, Hsu PK, Suo J, Lin TP. Pulmonary tuberculosis in the Taiwanese prison population. J Formos med Assoc. 2002; 101(8):537-41.

Conninx R, Eshaya-Chauvin B, Reyes H. Tuberculosis in Prisons. Lancet 1995; 346:238-39.

Conninx R, Pfyffer GE, Mathieu C, Savina D, Debacker M, Jafarov F, Jabrailov I, Ismailov A, Mirzoev F, de Haller R, Portaels F. Drug resistant tuberculosis in prisons in Azerbaijan: case study. BMJ 1998; 316(7142):1423-5.

Curtis R, Friedman SR, Neaigus A, Jose B, Goldstein A, Jose B, Goldstein M, Des Jarlais DC. Implications of directly observed therapy in tuberculosis control measures among IDUs. Public Health Rep. 1994; 109(3):319-27.

Daley CL, Small PM, Schecter, GF Schoolnik GK, McAdam RA, Jacobs WR Jr, Hopewell PC. An Outbreak of Tuberculosis with Accelerated Progression among Persons Infected with Human Immunodeficiency Virus: An Analysis Using Restriction Length Polymorphism. N Engl J Med 1992; 326:231-235.

Dara M, Grzemska M, Kimerling ME, Reyes H, Zagorsky A. Guidelines for control of tuberculosis in prisons. The Tuberculosis Coalition for Technical Assistance 2009.

Darbyshire JH. Tuberculosis in prisons. BMJ. 1989; 299(6704):874.

Drobniewski F. Tuberculosis in prisons - forgotten plague. Lancet 1995; 346(8980):948-9.

Drobniewski F, Taylor E, Ignatenko N, Paul J, Connolly M, Nye P, Lyagoshina T, Besse C. Tuberculosis in Siberia: 1. An epidemiological and microbiological assessment. Tuber Lung Dis 1996; 77(3):199-206.

Fernández-Martín JI, Fernández de la Hoz K, Catalán S, Alonso Sanz M, Chavese F. Med Clin (Barc). 2000; 115(7):246-50.

Hanau-Berçot B, Grémy I, Raskine L, Bizet J, Guttierrez MC, Boyer-Mariotte S, Brégeault A, Lagrange PH, Sanson Le Pors MJ. A one-year prospective study (1994-1995) for a first evaluation of tuberculosis transmission in French prisons. Int J Tuberc Lung Dis. 2000 Sep; 4(9):853-9.

Hutton. Results of a 29-state survey of tuberculosis in nursing homes and correctional facilities. Public Health Rep. 1993 May-June; 108(3):305-14.

Ijaz K, Yang Z, Templeton G, Stead W.W, Bates J.H, Cave M.D. Persistence of a strain of Mycobacterium tuberculosis in a prison system. Int J Tuberc Lung Dis 8(8):994-1000 c 2004 IUALTD.

Jittimanee SX, Ngamtrairai N, White MC, Jittimanee S. A prevalence survey for smear-positive tuberculosis in Thai prisons. Int J Tuberc Lung Dis. 2007; 11(5):556-61.

Kendig N. Tuberculosis control in prisons. Int J Tuberc Lung Dis. 1998 Sep;2(9suppl1):S57-63

Kimerling ME, Kluge H, Vezhnina N, Iacovazzi T, Demeulenaere T, Portaels F, Matthys F. Inadequacy of the current WHO retreatment regimen in a central Siberian prison: treatment failure and MDR-TB. Int J Tuberc Lung Dis 1999; 3(5):451-3.

Klopf. Tuberculosis control in the New York State Departement of Correctional Services: a case management approach. Am J Infect Control 1998; 26(5):534-7.

Koffi N, Ngom AK, Aka-Danguy E, Seka A, Akoto, Fadiga D. Smear positive pulmonary tuberculosis in a prison setting: experience in the penal camp of Bouaké, Ivory Coast.Int J Tuberc Lung Dis 1997; 1(3):250-3.

Lawn SD, Zumla AI. Tuberculosis. (www.thelancet.com Published online March 18, 2011 DOI:10.1016/S0140-6736(10)62173-3; consulted 31/12/2011).

Legrand J, Sanchez A, Le Pont F, Camacho L, Larouzé B. Modeling the impact of tuberculosis control strategies in highly endemic overcrowded prisons. PLoS One 2008 May 7; 3:e2100.

Lönnroth K, Jaramillo E, Williams BG, Dye C, Raviglione M. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. Soc Sci Med. 2009; 68(12):2240-6.Epub 2009 Apr 23.

Lönnroth K, Castro KG, Chakaya JM, Chauhan LS, Floyd K, Glaziou P, Raviglione MC. Tuberculosis control and elimination 2010-50: cure, care, and social development. Lancet 2010 May 22; 375(9728):1814-29.

Macallan DC. Nutrition and immunity: lessons from the past and new insights into the future. Amer J Clin Nutr 1991; 53:1087-1101.

MacIntyre CR, Kendig N, Kummer L, Birago S, Graham NM. Impact of tuberculosis control measures and crowding on the incidence of tuberculosis infection in Maryland prisons. Clin Infect Dis 1997; 24(6):1060-7.

MacNeil JR, Lobato MN, Moore M. An Unanswered Health Disparity: Tuberculosis among Correctional Inmates, 1993 through 2003. American Journal of Public Health, October 2005. Vol 95,N° 10.

Maher D, Grzemska M, Coninx R, Reyes H. Guidelines for the control in prisons. Geneva: WHO 1998 "(WHO/TB/98.250).

March F, Coll P, Guerrero RA, Busquets E, Caylà JA, Prats G. Predictors of tuberculosis transmission in prisons: an analysis using conventional and molecular methods. AIDS 2000; 14:525-35.

Martin V, Gonzalez P, Cayla JA, Mirabent J, Canellas J, Pina JM, Miret P. Case-finding of pulmonary tuberculosis on admission to a penitentiary center. Tuberc Lung Dis 1994; 75(1): 49–53.

Mazurek GH, Jereb J, Vernon A, LoBue P, Goldberg S, Castro K; IGRA Expert Committee; Centers for Disease Control and Prevention (CDC). Updated guidelines for using

Interferon Gamma Release Assays to detect Mycobacterium tuberculosis infection - United States, 2010. MMWR 2010 Jun 25; 59(RR-5):1-25.

Murray M, Oxlade O, Linh HH. Modeling social, environmental and biological determinants of tuberculosis. Int J Tuberc Lung Dis 2011; 15:S64-S70.

Nardell EA. Tuberculosis in homeless, residential care facilities, prisons, nursing homes, and other close communities. Semin Resp Infect 1989; 40(3):206-15.

Noeske J, Kuaban C, Amougou G, Piubello A, Pouillot R. Pulmonary tuberculosis in the Central Prison of Douala, Cameroon . East Afr Med J. 2006; 83(1):25-30

Noeske J, Ndi N, Mbondi S. Controlling tuberculosis in prisons against confinement conditions: a lost case? Experience from Cameroun. Int J Tuberc Lung Dis 2011; 15(2):223-7.

Nyangulu DS, Harries AD, Kang'ombe C, Yadidi AE, Chokani K, Cullinan T, Maher D, Nunn P, Salaniponi FM. Tuberculosis in a prison population in Malawi. Lancet 1997; 350(9087):1284-7.

O'Grady J, Hoelscher M, Atun R, Bates M, Mwaba P, Kapata N, Ferrara G, Maeurer M, Zumla A. Tuberculosis in prisons in sub-Saharan Africa--the need for improved health services, surveillance and control. Tuberculosis (Edinb). 2011; 91(2):173-8.

Pelletier AR, DiFerdinando GT Jr, Greenberg AJ, Sosin DM, Jones WD Jr, Bloch AB, Woodley CL. Tuberculosis in a correctional facility. Arch Intern Med. 1993; 153(23):2692-5.

Rao NA. Prevalence of pulmonary tuberculosis in Karachi central prison. J Pak Med Assoc 2004; 54(8):413-5.

Rieder HL. Epidemiological Basis of Tuberculosis Control. Paris 1999 (IUATLD).

Rieder HL, Anderson C, Dara M, Hauer B, Helbling P, Kam KM, Zwahlen M. Methodological issues in quantifying the magnitude of the tuberculosis problem in a prison population. Int J Tuberc Lung Dis 2011; 15:662-667.

Rutta E, Mutasingwa D, Ngallaba S, Mwansasu A. Tuberculosis in a prison population in Mwanza, Tanzania (1994-1997). Int J Tuberc Lung Dis 2001; 5(8):703-6.

Sanchez A, Gerhardt G, Natal S, Capone D, Espinola A, Costa W, Pires J, Barreto A, Biondi E, Larouzé B. Prevalence of pulmonary tuberculosis and comparative evaluation of screening strategies in a Brazilian prison. Int J Tuberc Lung Dis 1995; 9(6):589.

Valway SE, Greifinger RB, Papania M, Kilburn JO, Woodley C, Diferdinando GT, Dooley SW. Multidrug-resistant tuberculosis in the New York State prison system, 1990-1991. 1993.

Sretrirutchai S, Silapapojakul K, Palittapongarnpim P, Phongdara A, Vuddhakul V. Tuberculosis in Thai prisons: magnitude, transmission and drug susceptibility. Int J Tuberc Lung Dis 2002; 6(3):208-14.

Stead WW. Undetected tuberculosis in prison. Source of infection for community at large. JAMA. 1978; 240(23):2544-7.

Stuckler D, Basu S, McKee M, King L. Mass incarceration can explain population increases in TB and multidrug-resistant TB in European and central Asian countries. Proc Natl Acad Sci USA 2008; 105(36):13280-5.

Uys P, Marais BJ, Johnston-Robertson S, Hargrove J, Wood R. Transmission elasticity in communities hyperendemic for tuberculosis. Clin Infect Dis 2011; 52(12):1399-404.

Valway SE, Richards SB, Kovakovich J, Greifinger RB, Crawford JT, Dooley SW. Outbreak of multi-drug-resistant tuberculosis in a New York State prison, 1991. Am J Epidemiol 1994; 140(2):113-22.

Wong MY, Leung CC, Tam CM, Kam KM Ma CH, Au KF. TB surveillance in correctional institutions in Hong Kong 1999-2005. Int J Tuberc Lung Dis 2008; 12(1):93-8.

World Health Organization. Regional Office for Europe. Status Paper on Prisons and Tuberculosis. Copenhagen 2007 (EUR/07/5063925) (WHO 2007a).

World Health Organization. Health in prisons. A WHO guide to the essentials in prison health. Copenhagen 2007 (EUR/07/5063912) (WHO 2007b).

World Health Organization. The Stop TB Strategy
(http://www.who.int/tb/stategy/stop_tb_strategy/en.index.html).
Accessed 15.12.2011)(WHO 2010a).

Stop TB partnership. Global plan to stop TB 2011-2015: Transforming the Fight – Towards Elimination of Tuberculosis.
(http://www.stoptb.org/assets/documents/global/plan/TB_GlobalPlanToStopTB 2011-2015.pdf).
Accessed 01.12.2011.

World Health Organization. Use of tuberculosis interferon gamma release assays (IGRAs) in low- and middle-income countries: policy statement. WHO/HTM/TB/2011.18. Geneva: WHO 2011. (WHO 2010b).

21

Social Determinants of Health in Deaf Communities

Scott R. Smith and Nancy P. Chin
Department of Community and Preventive Medicine, University of Rochester
USA

1. Introduction

The social determinants of health among deaf people are similar to those of hearing people in the United States. The built environment, neighborhood cohesiveness, access to health care, education, and affordable housing all shape the health of individual deaf people as they do hearing people. However, in addition to these similarities, deaf people face a unique combination of social and communication barriers which appear to have resulted in health disparities between hearing and deaf people. These barriers often are more pervasive for deaf people than hearing people and have important public health implications for deaf communities. The ultimate goal of this chapter is twofold: 1) to discuss social factors unique to deaf people which might contribute to health disparities between hearing and deaf people; and 2) to suggest programmatic and systematic approaches to help close these health gaps.

2. Pertinent socioecological barriers and issues affecting deaf people and their well-being

In 2006, the Centers for Disease Control-funded National Center for Deaf Health Research (NCDHR) at the University of Rochester partnered with the local deaf community to develop and implement a unique Deaf Health Survey (DHS) in American Sign Language (ASL). The survey findings indicated three important health-related disparities between the deaf people who took the survey and the hearing people who live in the same county including: 1) increased rates of cardiovascular risks including obesity; 2) increased risks for intimate partner violence; and 3) increased reports of suicidal attempts (Barnett et al., 2011). These findings are consistent with other previous studies indicating that deaf people appear to be experiencing significant disparities in their health and mental health literacy and outcomes (Margellos-Anast et al., 2006, Maxwell-McCaw, 2001; Hindley et al., 1994). Whereas these surveys and studies have helped to identify specific health disparities that deaf people experience, they did not include items that explored the reasons why deaf people might be experiencing these health disparities.

Preventing such health disparities in deaf people requires understanding the factors that contribute to their health inequities. In the first part of this chapter, a four-level socioecological perspective is used to describe the pertinent challenges facing deaf people's public health to better appreciate the complex interplay between these 1) individual, 2) interpersonal (relationship), 3) community, and 4) societal barriers (Dahlberg & Krug, 2002). In the second portion of this chapter, potential prevention strategies encompassing a continuum of activities that address the multiple levels of the socioecological model are proposed in hopes that these approaches might be more likely to sustain prevention efforts over time than any individual intervention.

2.1 Individual: The impact of deafness on the person

Because there is no single widely-accepted definition of what makes a person "deaf" as compared to a person who might have some level of hearing loss, this chapter, for the sake of simplicity, will focus on people who have hearing loss that is significant enough that they are unable to have unfettered spoken conversations with other people; and that this hearing loss occurred at birth or within the first three years of life. Using this definition of deafness, approximately 1-2 in 1000 people worldwide, regardless of gender or race/ethnicity variations, are deaf (Watkin & Baldwin, 2011) with about 4.8 million Americans in 1994 who reported being unable to hear or understand speech (Ries, 1994). As elaborated in the remainder of this chapter, because the world is an auditory-dominated realm of hearing people, deafness has significant implications for an individual's well-being on all socioecological levels including their interpersonal relationships with others close to them, with their communities, and with the society at large.

Early-onset deafness cannot be perceived as being equal to late-onset deafness because the individual who becomes deaf within the first few years of life experiences life differently than an individual who might lose her hearing later in life. An individual who loses her ability to hear normally after 3 years of age or later in life is much more likely to view her hearing loss as a true disability than an individual who has never had typical hearing. This subtle, but important, distinction is the first barrier that faces a deaf child in her interactions with the auditory-dominant world. The hearing world automatically views the deaf child as being disabled because the child does not have a functional ability that most hearing people could not imagine living without. As a result, the first focus is on remediating the deaf child's audiological problems.

Virtuous efforts by well-meaning hearing people and professionals to remediate deaf children's audiological problems exert some unintended consequences on the deaf child. Based on their own experiences, many hearing people and professionals believe that being able to hear and speak sounds, even at a rudimentary level, is essential for effective language development and social functioning. As they have themselves never experienced the need to do so, most hearing people might not realize or appreciate the powerful utility of using intact vision to teach deaf children language and about the world around them. For example, many hearing people do not realize that deaf people can be taught the phonological rules necessary for strong reading and writing skills through sign language without any need for the deaf person to be able to hear any sounds. As a result, many deaf children are trained to master language and social functioning using a repaired and often less-than-optimal information channel (auditory) when another intact information channel (vision) is readily available. Consequently, the natural "path of least resistance" through sufficient visual information for language and social development is denied for many deaf children.

Denying young children their paths of least resistance to language and social development has important implications for any children's cognitive and psychosocial development. Despite best efforts to date, too many deaf people continue to struggle with their mastery of English in the United States. The most recent available data in the United States indicate that deaf high school students are graduating with a median of English reading skills comparable to hearing fourth graders (Holt, 1993). Although anecdotal evidence suggest that deaf people might be improving their reading skills with recent improvement in deaf education, it is still unlikely that, even in 2011, most deaf people are reading English at the same level as most hearing people. Furthermore, as deaf children are forced to work through their developmental challenges using less-than-optimal information channels, deaf children are made to realize that they are different from other people at a very early age. Deaf children are forced to adjust and adapt to their environments that do not promote their easiest path to development as a person. A consequence of the different ways that deaf people have adjusted and adapted to their environments can be observed in the various labels that deaf people use to describe themselves. Some deaf people describe themselves as being hearing-impaired or hard-of-hearing while other deaf people identify themselves as being Deaf with capital 'D' to emphasize their cultural identification and some others who call themselves deaf without the capital 'D'. Under certain circumstances, some deaf people might also attempt to completely deny their hearing loss.

2.2 Interpersonal: The deaf person is usually a visual minority within their own auditory-dominated family

Unlike most people of minority groups who are usually born to parents who are also members of the same minority group, approximately 96% of deaf people are born to hearing parents (Moores, 2001), who usually do not know very much about deafness and deaf people. As a result, most deaf children and adults are minorities within their own families from the day that they become deaf. Often in an attempt to be inclusive, anxious hearing parents of deaf children might try to 'fix' their child by helping him or her to hear and speak as best as possible. To that end, these parents typically enlist a cadre of professional audiologists, speech therapists, and educators to stimulate English language development and communication in their deaf children without recognizing or realizing the strengths of visual learning strategies that might be much more effective for many deaf children and people (Meadow-Orlans, 1994). Such auditory-based strategies tend to focus on augmenting and boosting deaf children's residual hearing through technologies such as hearing aids and cochlear implants accompanied with intensive speech and hearing training without much attention to deaf children's visual strengths.

As many scientific studies of popular auditory-based interventions have focused on auditory-based outcomes such as improved speech and sound recognition (Waltzman & Roland, 2005; Yoshinaga-Itano et al., 2000; Yoshinaga-Itano et al., 1998), such intervention strategies coupled with cochlear implantations and universal newborn hearing loss screenings have become widely accepted as the gold standard treatment for deafness in children (Brentari, 2010; Snow & Wackym, 2008; Christensen & Leigh, 2002). In the United States and other developed nations, the diffusion of such intervention strategies including a surgical procedure has been very rapid and widespread despite the fact that the success rates of these interventions tend to vary greatly depending on many other factors such as parental educational levels with failure rates approaching 20-50% in some studies (Gulya et al., 2010; Peterson et al., 2010; Bouchard et al., 2008). Studies that examine other measurable and meaningful outcomes of these auditory-based interventions for deaf children such as essential language development and long-term quality of life and studies that include alternative strategies to develop deaf children's language through visual modalities such as signed languages (Kushalnagar et al., 2011, Patrick et al., 2011; Punch & Hyde, 2011) have only very recently begun to penetrate the medical literature that influence hearing professionals and parents in their decision-making regarding the best ways to promote deaf children's and deaf people's long-term outcomes.

Many deaf children and adults cannot fully communicate with their hearing parents and family members. Deaf children are usually not able to begin communicating with their hearing parents until much later when they are diagnosed and provided with interventions. Furthermore, even after diagnosis and treatment, many deaf children are unable to experience unfettered communication with their hearing parents and family members because their parents and health care providers might not realize that their deaf children also need enhanced visual information (Zimmerman, 2009; Krausneker, 2008; Lane, 2005). A common criticism that many older deaf children and adults have about their hearing parents and families relate to the enormous amount of time and energy spent in speech therapy, learning how to use their voice and read lips, rather than working on normal childhood developmental tasks including learning a language and learning how to read and write. These factors have important implications in the psychological development of deaf children's relationships with their hearing parents. Many deaf children describe relating to their hearing parents in ways that are different than hearing children of hearing parents (Preston, 1995).

Because they are unable to easily communicate with their hearing family members and participate in family activities, some deaf people develop a significant mistrust of their hearing parents. Based on Erikson's psychological development paradigm (Erikson, 1950), deaf children's mistrust of their hearing parents can interfere with deaf children's psychological development. Sometimes, their mistrust evolves into more significant resentment of hearing people in general (Hauser et al., 2010). Such "deaf-ism" might lead deaf people to prefer communicating solely with other like-minded deaf people and completely resist any messages from any hearing people, regardless of intent or context. This is likely to affect how these deaf people choose to interact with and receive messages from systems that are run by majority hearing people such as the health care and public health systems.

2.3 Community: The deaf person is usually a visual minority within their own auditory-dominated community

Even deaf people who experience the best possible hearing through hearing aids and/or cochlear implants are not always able to completely access all the auditory information that might be available in their community environments such as overhearing a nearby conversation in the public. As a result, almost all deaf people do not have the same “fund of information” – or amount of general knowledge - as hearing people who might have comparable education (Pollard, 1998). Even when well-educated, deaf people’s inability to overhear tremendous amounts of information that takes place in their environment limits the amount of knowledge that they are able to acquire over time. For example, a typical deaf person would not benefit from listening to other people having a conversation on the bus about the spread of H1N1 virus, listening to their family members discussing grandfather’s diabetes, or listening to their co-workers having conversations in the hallway about a workplace wellness program. When extrapolated to an entire lifetime of missed opportunities for further information, it becomes clear that almost all deaf people experience restricted informal learning and they are underexposed to information in general. Unless these information gaps are acknowledged, hearing people might attempt to present deaf people with information and messages that take for granted a certain amount of “common” knowledge. For example, a hearing doctor might assume that her deaf patient understands what “cholesterol” is while carefully discussing treatment options without realizing that less than 10% of young deaf people actually know what cholesterol is (Smith et al., manuscript in preparation).

Aside from auditory information that is available in community environments, deaf people’s limited English literacy has important implications for public health as most health information available in written forms such as brochures, flyers, signs, newspapers, magazines, captioned TV, and Internet are typically written for people who are able to read English at the seventh grade level or higher (Freda, 2005). Recent focus groups of young deaf people in Rochester, NY confirm this trend with many young deaf people describing significant difficulties understanding health information (Smith et al., manuscript in preparation).

Another important implication of the fact that most deaf people are underexposed to information is that they are also being underexposed to stimuli that might prompt them to seek out more information and improve their health knowledge and literacy. For example, a hearing person who might overhear an interesting conversation about someone who had a heart attack at a young age might be then prompted to seek out more information about how she could reduce her own risks for having a heart attack. A deaf person in the same scenario most likely would not hear or understand this conversation and therefore, she would not be prompted to seek out more information to reduce her risks for a heart attack. This has important implications for the future where health care consumers are expected to take on more responsibilities for their own health including being able to seek out and understand health information from various sources.

2.4 Societal: At the employment level, the deaf person often loses out to the “bottom line” philosophy of most businesses and organizations

Deaf people’s inability to fully access the same amount of information and knowledge from their environments as hearing people also has important implications for their employment opportunities and performance. In the United States, recent changes in education legislation have allowed many deaf children and people to attain educational levels that are approaching those of hearing children and people. By law, deaf children are entitled to specialized support services and various accommodations such as sign language interpreters, speech-language therapy, auditory training, specialized teachers for the deaf, teaching methods that incorporate increased visual information, and specialized supports for English reading and writing skills. However, despite the recent passage of the Americans with Disabilities Act (ADA) in the United States, these deaf people's educational achievements have not always translated into employment successes because the ADA places the burden of the costs of services and accommodations that might have afforded a deaf person her educational success on the shoulders of employers. As a result, employers might hesitate to hire, retain, or promote deaf employees because of the undesired impact on their bottom line. Unfortunately, ADA also places the burden of requesting such expensive accommodations and undesired costs on deaf people who are expected to negotiate such issues directly with the people who they are trying to convince to hire them despite their hearing issues. As a result, some deaf people attempt employment without appropriate levels of accommodations and/or supports that would allow them to perform well in their jobs. A recent survey of deaf adults in Rochester, NY confirms the unfortunate socioeconomics of deaf people. Deaf people of all educational levels report earning incomes that are significantly less than hearing people of comparable educational levels (Barnett et al., 2011).

The implications of deaf people's lower incomes, despite their formal education, cannot be overstated and is most likely related to their lifelong restricted informal education. Aside from the usual consequences of lower socioeconomic status on health care access, health insurance coverage issues, transportation, and many other barriers, deaf people who realize that they are being underpaid for their education and training might be prone to some significant psychosocial difficulties and employment underperformance.

On a societal level, employment anti-discriminatory legislations and regulations such as the ADA have proven to be effective in promoting employment of many minorities and people with many types of disabilities, but they do not sufficiently address the employment needs of many deaf people. This critical structural barrier to deaf people's optimal health could be effectively addressed and eliminated with some appropriate public policy changes.

2.5 The big picture: Negative synergetic effects of combination of multi-leveled barriers

The combined and cumulative effect of the various barriers that deaf people experience across multiple levels of the socioecological model appear to be more complicated than the effect of these barriers on hearing people and they are likely contributing to the poorer health of deaf people (see Figure 1). Unfortunately, because of these barriers, deaf people's unique challenges have been invisible to and, usually unintentionally, neglected by health care research and public health policy.

3. Potential solutions

3.1 Use the cultural model, not the medical model of deafness

An important consideration that must be accounted for in any serious public health endeavors involving deaf people is the fact that many deaf people, especially those who use sign language, do not accept the notion that their deafness is truly a disability.

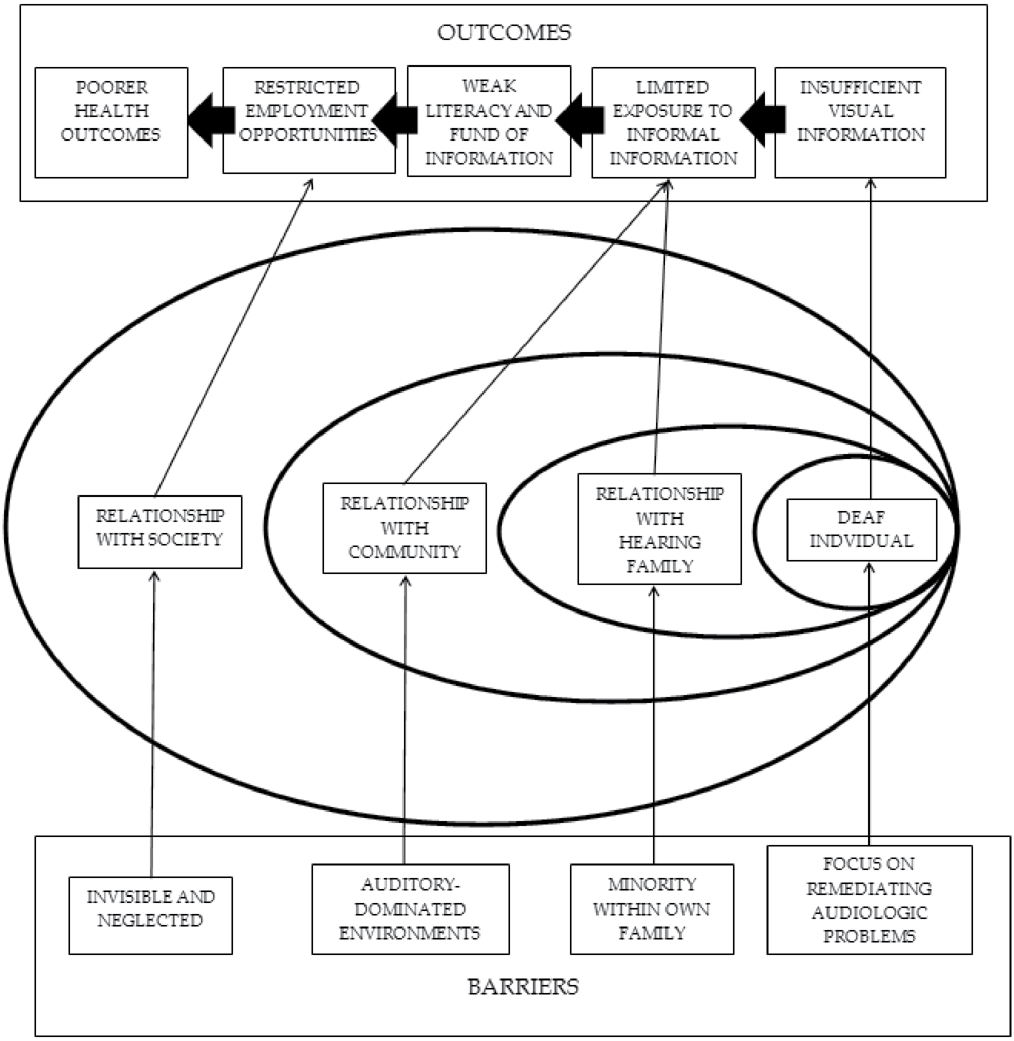


Fig. 1. A Socioecological Perspective of Deaf People's Barriers to Optimal Health

Most deaf adults are able to lead reasonably independent lives as people who use their normally functioning eyes to drive themselves to a well-earned family vacation in Florida and communicate easily with others through sign language. They often perceive the scientific and medical worlds as being misguided in their attempts to cure hearing losses, especially in infants and young children who are not capable of making such an important life choice. Regardless of their communication modalities or use of auditory-based means, most deaf people are inherently more visual-oriented than hearing people. Visual information is processed differently in the human brain than auditory information. Deaf people who use sign language tend to perceive themselves as having valuable perspectives and understandings of the world that are different than those of most hearing people. As a result, they often perceive their deafness as actually conferring them with a unique persona and various advantages that might outweigh all the disadvantages they experience daily.

The fact is many deaf people who use sign language do not perceive their deafness to be a disability. This has important implications for public health endeavors that target this particular population. Serious attempts to conduct health research and design interventions with these deaf people require interested researchers to abandon the usual scientific and medical paradigm that assumes all deaf people would rather to hear no matter what. NCDHR has been able to develop successful working collaborations with the deaf communities in Rochester, NY by making a conscious decision to pursue a cultural framework where the focus is on the health and well-being of deaf people, regardless of their background or communication preferences. Their focus is not on the auditory problems of deaf people. Using this cultural framework has enabled the process of engaging deaf people in health research to begin in Rochester, NY, which has clearly helped many local deaf people to become more aware of their health issues.

3.2 Use equitable partnerships to include deaf people in health research

Because of deaf people's recognized inclinations to resist engaging with the health care and public health systems, the NCDHR utilized a community-based participatory approach (Barnett et al., 2011) to gain the trust of deaf people and collaborate with them for synergistic results. Community-Based Participatory Research (CBPR) approach differs from traditional research in many ways. CBPR encompasses an equitable partnership between researchers and community members, ideally from the very beginning of the research process (Wallerstein & Duran, 2003). NCDHR engaged the Rochester deaf communities by sharing power, resources, credit, results, and knowledge at each stage of research. For perhaps the very first time, deaf people who use sign language were given an essential role in their own health research. Thereby, their involvement is enhanced by the increased stakes that they possess in the research process and the first step towards trust is made.

As a result of NCDHR's ongoing collaboration with the Rochester deaf communities, specific survey and data collection methods have been developed to make health-related measures accessible to many more deaf people than typical health surveys. Based on these experiences, Barnett and others propose that health surveys that might include deaf people should ask individuals who report a hearing loss a set of questions that would provide more meaningful information about deaf people and their health (Barnett et al., 2011). These items include: 1) degree of hearing loss (mild, moderate, severe, profound); 2) laterality of their hearing loss (bilateral or unilateral); 3) age-at-onset of deafness (birth, before age 1, before age 3, or after age 3); 4) presence of other deaf family members (yes, one; yes, more than one; no); 5) preferred communication modality (sign language, spoken language, written language, or some combination); 6) perceived quality of communication with parents (good, average, poor, no communication); 7) educational setting (school for the deaf, school for hearing students); and 8) self-reported functional literacy such as how well can you read English? (good, average, poor, unable to read). For example, it would be irrational to expect a 50-year woman who has a mild unilateral hearing loss that began a few years ago to have the same communication, language, socioeconomic, and public health issues as a 50-year-old woman who has bilateral profound hearing loss from birth, attended a school for deaf, uses sign language, doesn't communicate with her parents at all, has 2 deaf siblings, and self-reports poor English reading skills. Unfortunately, to date, very few regional and national health care surveys include such questions and as a result, data that might help to explain deaf people's health disparities are limited.

3.3 Create and implement health educational materials that are truly accessible to deaf people

An enormous amount of health education materials that encompass critical personal and public health information needs to be developed and distributed to deaf people and communities. Such information might include what foods to eat to avoid a heart attack and what supports might be available to a victim of domestic violence. Health information that is accessible to a wide spectrum of deaf people will help overcome deaf people's health knowledge deficits and health literacy challenges.

Most deaf people are visually-oriented regardless of their hearing abilities. Health information developed for a deaf audience must include significant amounts of concrete and tangible visual information to convey the basic message. When possible, all health information should be presented in sign language along with written health information presented at the lowest reading level feasible to communicate the meaning of the message. Effective health information for deaf people will also need to include further explanations and clarifications of certain terminologies that might be expected to be common knowledge such as "cholesterol" to overcome their "fund of information" deficits. Visual conversational exchanges of information between two or more people also appear to be effective in conveying complicated concepts for some deaf people. Including and employing deaf professionals and community members to develop and distribute this type of information ensures optimal utility. Recent technological advances in mobile communication devices hold much promise for social networking capabilities as a tool for information dissemination. These modified health educational materials are likely to be valuable for at least some hearing people as well.

3.4 Educate and employ deaf people in the public health world

Educating, training, hiring, accommodating, mentoring, and promoting deaf professionals in the public health system would promote engagement of deaf people in the research process as researchers. The University of Rochester has pursued several endeavors to attain this goal including forming an educational pipeline to link high schools that serve deaf students, colleges and universities that educate deaf students, and places of employment that employ deaf professionals. For example, a deaf student from a local school for the deaf recently successfully participated in the University of Rochester's Science and Technology Entry Program (STEP) for members of underrepresented minority high school students. She is now a very successful college student at the local university pursuing a science major. The University of Rochester has also recently hired two full-time deaf faculty members and the NCDHR employs several full-time and part-time deaf employees along with a full-time staff sign language interpreter and many other accommodations including real-time captions for meetings.

3.5 Modify employment anti-discriminatory policies to better meet deaf people's employment needs

On a broader policy level in the United States, deaf people who use sign language would benefit from modifications in the current employment anti-discriminatory legislation to fully eradicate the remaining barriers to employment. Ideally, the Americans with Disabilities Act (ADA) would be modified to recognize that deaf people, unlike many other disabilities, require accommodations and services that are ongoing and too expensive to expect employers to shoulder the costs. In the case of deaf Americans, the ADA needs to be modified to become a funded mandate where the costs of accommodations and services to employ deaf people would be covered by the federal government to level the employment playing field for deaf people. Such a change would also eliminate the burden that ADA places on deaf employees to negotiate, compromise, and even beg for accommodations and services that would increase their employment success. Whereas such a program might require a significant financial investment, the potential long-term cost-savings of employing more deaf people and helping to ensure their health has yet to be realized. Between 30% and 60% of deaf Americans continuing to be unemployed (US Census Bureau, 2002) and many more being underemployed, as evidenced by their lower incomes despite their education (Barnett et al., 2011).

4. Conclusions and implications: Deaf and hearing people must work together for the best interests of deaf people

Working with deaf communities presents the public health systems with challenges that are both similar and more complicated than those faced by members of gender, racial/ethnicity, and linguistic minorities. However, through community-based participatory research approaches, groundbreaking gains have been achieved with growing and evolving collaborations between the University of Rochester and the deaf communities of the local region. Such gains and culminated progress over several years hold the promise for even more successful achievements in improving the health and well-being of a group of people who have been historically misunderstood and neglected by public health systems. Effective dissemination and long-term successful implementation of this information into the medical and public health world needs much support from open-minded hearing people who are willing to listen not only with their ears, but with their eyes and hearts.

5. References

Barnett S, McKee M, Smith SR, & Pearson TA. (2011). Deaf sign language users, health inequities, and public health: opportunity for social justice. Preventing Chronic Disease, Vol. 8, No. 2, pp. A45.

Barnett, S., Klein, J.D., Pollard, R.Q., Samar, V.J., Schlehofer, D., Starr, M.J., Sutter, E., Yang, H., & Pearson T.A. (2011). Community participatory research to identify health inequities with deaf sign language users. American Journal of Public Health, (in press).

Bouchard, M.E., Ouellet, C., & Cohen, H. (2008). Speech development in prelingually deaf children with cochlear implants. Language and Linguistics Compass, Vol.2, pp. 1-18.

Brentari, D. (Ed.). (2010). Sign Languages, Cambridge University Press, Cambridge, UK.

Christensen, J. & Leigh, I. (2002). Cochlear Implants in Children: Ethics and Choices, Gallaudet University Press, Washington, DC.

Dahlberg, L.L. & Krug, E.G. (2002). Violence-a global public health problem. In: World Report on Violence and Health. Geneva, Switzerland: World Health Organization, E. Krug, L. Dahlberg, J. Mercy, A. Zwi, & R. Lozano (Eds.), pp. 1–56.

Erikson, E.H. (1950). Childhood and society. Norton, New York (1950).

Freda, M.C. (2005). The readability of American Academy of Pediatrics patient education brochures. Journal of Pediatric Health Care. Vol. 19, No. 3, pp.151-6.

Gulya, A.; Minor, L., & Poe, D. (2010). Glasscock-Shambaugh's Surgery of the Ear, (6th edition), People's Medical Publishing House USA, Shelton, CT.

Hauser, P.C., O'Hearn, A., McKee, M., Steider, A., & Thew, D. (2010). Deaf epistemology: deafhood and deafness. American Annals of the Deaf, Vol. 154, No. 5, pp. 486-96.

Hindley, P., Hill, P.D., McGuigan, S., & Kitson, N. (1994). Psychiatric disorder in deaf and hearing impaired children and young people: a prevalence study. Journal of Psychiatry and Pscyhology, Vol. 35, pp. 917-34.

Holt, J. (1993). Stanford Achievement Test – 8th edition: Reading comprehension subgroup results. American Annals of the Deaf, Vol. 123, pp. 172-5.

Krausneker, V. (2008). The protection and promotion of sign languages and the rights of their users in Council of Europe member states: needs analysis, In: Council of Europe – Partial Agreement in the Social and Public Health Field, 15.07, 11, Available from http://www.coe.int/t/DG3/Disability/Source/Report_Sign_languages_ final.pdf

Kushalnagar, P, Topolski, T.D., Schick, B., Edwards, T.C., Skalicky, A.M., Patrick, D.L. (2011). Mode of communication, perceived level of understanding, and perceived quality of life in youth who are deaf or hard of hearing. Journal of Deaf Studies and Deaf Education, Vol. 16, No.4, pp. 512-23.

Lane, H. (2005). Ethnicity, ethics, and the Deaf-world. Journal of Deaf Studies and Deaf Education, Vol.10, No.3, pp. 291–310.

Margellos-Anast, H., Estarziau, M., Kaufman G. (2006). Cardiovascular disease knowledge among culturally deaf patients in Chicago. Preventive Medicine, Vol. 42, No. 3, pp. 235-9.

Marschark, M., Rhoten, C., & Fabich, M. (2007). Effects of cochlear implants on children's reading and academic achievement. Journal of Deaf Studies and Deaf Education, Vol.12, No.3, pp. 269-282.

Maxwell-McCaw, D. (2001). Acculturation and Psychological Well-Being in Deaf and Hard of Hearing People. Unpublished doctoral dissertation, George Washington University, Washington DC.

Meadow-Orlans, K. (1994). Stress, social support, and deafness: perceptions of infants' mothers and fathers. Journal of Early Intervention, Vol.18, No.1, pp. 91–102.

Moores, D. (2001). Educating the Deaf: Psychology, Principles, and Practice, Houghton Mifflin, Boston.

Patrick, D.L., Edwards, T.C., Skalicky, A.M., Schick, B., Topolski, T.D., Kushalnagar, P., Leng, M., O'Neill-Kemp, A.M., & Sie, K.S. (2011). Validation of quality-of-life measure for deaf or hard of hearing youth. Otolaryngology Head Neck Surgery, Vol. 145, No. 1, pp. 137-45.

Peterson, N., Pisoni, D., & Miyamoto, R. (2010). Cochlear implants and spoken language processing abilities: review and assessment of the literature. Restorative Neurology and Neuroscience, Vol.28, No.2, pp. 237-250.

Pollard, R.Q. (1998). Psychopathology. In: Psychological Perspectives on Deafness, Volume 2, M. Marschark & D. Clark (Eds.), pp. 171-97, Lawrence Erlbaum, Inc., Mahwah, NJ.

Preston, P. (1995). Mother father deaf: the heritage of difference. Social Science and Medicine. Vol. 40, No. 11, pp. 1461-7.

Punch, R. & Hyde, M. (2011). Social participation of children and adolescents with cochlear implants: a qualitative analysis of parent, teacher, and child Interviews. Journal of Deaf Studies and Deaf Education, Vol.16, No.4, pp. 474-493.

Ries, P.W. (1994). Prevalence and Characteristics of Persons with Hearing Trouble: United States, 1990-1. National Center for Health Statistics. Vital Health Statistics, Section 10, pp. 188.

Smith, S.R. & Chin, N.P. (in preparation). Young deaf people's perspectives of their cardiovascular health knowledge and literacy.

Snow, J., Wackym, P. (2008). Ballenger's Otorhinolaryngology, (17th edition), People's Medical Publishing House, Shelton, CT.

U.S. Census Bureau. Americans with Disabilities: 2002 Household Economic Studies. U.S. Department of Commerce Economics and Statistics Administration. Current population reports by Erika Steinmetz.

Waltzman, S. & Roland, T. (2005). Cochlear implantation in children younger than 12 months. Pediatrics, Vol.116, No.4, pp. e487-e493.

Wallerstein, N. & Duran, B. (2003). The Conceptual, Historical, and Practical Roots of Community Based Participatory Research and Related Participatory Traditions. In: Community Participatory Research for Health, M. Minkler & N. Wallerstein (Eds.), pp. 27-52, Jossey Bass, San Francisco.

Watkin, P.M. & Baldwin, M. (2011). Identifying deafness in early childhood: requirements after the newborn hearing screen. Archives of Disease of Childhood, Vol. 96, No. 1, pp.62-6.

Yoshinaga-Itano, C., Sedey, A., Coulter, D., & Mehl, A. (1998). Language of early- and later-identified children with hearing loss. Pediatrics, Vol.102, No.5, pp. 1161 –1171.

Yoshinaga-Itano, C.; Coulter, D., & Thomson, V. (2000). The Colorado Hearing Screening Program: effects on speech and language for children with hearing loss. Journal of Perinatology, Vol.20, No.8s, pp. s132 –s142.

Zimmerman, A. (2009). Do you hear the people sing? Balancing parental authority and a child's right to thrive: the cochlear implant debate. Journal of Health and Biomedical Law, Vol.5, No.2, p. 309.

22

Anxiety and Emotional Discomfort in the School Environment: The Interplay of School Processes, Learning Strategies, and Children's Mental Health

L. Tramonte and J. D. Willms
University of New Brunswick
Canada

1. Introduction

Children and adolescents try to look strong and beautiful to their peers, deal with the pressure to succeed in sports, achieve good grades, and develop positive and lasting relationships. School is the place where children spend most of their daily hours, trying to meet these challenges while they define their identities. Consequently, experiences at school influence every aspect of development during childhood and adolescence, ranging from the construction of their intellectual capital to their emotional and physical well-being to the establishment of peer and child-adult relations (Eccles & Roeser, 2011).

Recently, researchers and educators have directed more attention to the relationship between the quality of learning environments – particularly effective teaching – and problems experienced by students in middle and high schools – such as disengagement, dissatisfaction with their schooling experience, and dropping out (Bowlby & McMullen, 2003; Pope, 2001; Willms, 2003). Relatively little attention has been paid to the role of schools in the identification of mental health problems or in helping to alleviate these problems, although the consequences of mental health problems on school attendance and achievement are palpable. As an example, the 2003 U.S. National Survey of Children's Health assessed emotional, cognitive, and behavioural problems in over 100,000 children and youth 0 to 17 years of age and found that children with chronic emotional, behavioural, and developmental problems that persisted for at least a year missed more than 10 days per year at school, three times that of their peers without these problems (Blanchard et al., 2006).

In most countries studied in the Program for International Students Assessment (PISA), there was considerable variation among schools in students' academic achievement and sense of belonging at school. Some of this variation is attributable to measureable aspects of classroom and school climate and structural features of the school system (Willms, 2003). However, PISA and other large-scale international studies do not consider variation in mental health outcomes such as anxiety and depression. It is not only large scale assessments which miss the mark of monitoring mental health and emotional well-being of the students; even national educational reforms also tend to ignore, or discard, this matter. In the United States, for instance, the No Child Left Behind Act (NCLB) placed increased pressure on schools to improve the academic achievement of students by requiring schools to track students' standardized achievement. At the same time, the budgets for preventive health services in schools were cut, and attention to symptoms related to poor emotional well-being, such as experiences of depression and anxiety, was neglected. Although there are legislative provisions that focus specifically on externalizing behaviors, such as the Gun-Free Schools Act, the Pro-Children Act, and the Safe and Drug Free Schools and Communities Acts, efforts to improve emotional well-being have received far less attention (Sznitman et al., 2011). The lack of attention to emotional well-being in the education policy discussion is particularly problematic in relation to growing evidence indicating that many early stressors can interfere with school achievement, compromise school completion, and lead to adverse mental health outcomes later in life, including chronic mental illness.

Depression is now the leading cause of disability in the world (Murray & Lopez, 1996); research on anxiety and depression in the school environment may be essential if we are to end what some have called the depression pandemic (Seligman et al., 2001). In fact, about one-half of adult mental health issues start before the mid-teens, but often treatment does not occur or is delayed until well into adulthood (Kessler et al., 2005; Wang et al., 2005). The number of children experiencing mental health problems is significant; results based on Canada's National Longitudinal Survey of Children and Youth (NLSCY) indicate that the prevalence of children and youth with anxiety problems ranges from 2% at age 2 to 12% at age 11, based on parents' assessments, and is higher prevalence for children aged 10 to 15, based on youth's self-assessments (Bagnell et al., 2009). There is also considerable variation in the prevalence of depression, ranging from 2% to 8% across the age range from 12 to 21, based on youth's self assessments (Willms, 2008). Earlier studies suggested that the prevalence of mental health problems for children and youth of this age is about 15% (Waddell et al., 2002). With some developmental differences, childhood depression is manifested in much the same fashion as adult variants; it impacts every facet of psychosocial functioning, including the family system, parent-child relationships, peer relationships, and school functioning. In extreme cases it can lead to suicide (Stark & Smith, 1995).

The literature offers a wide range of reported prevalence of behavioural disorders in children and youth, with estimates typically ranging from 10% to 20%. In a review by Roberts et al. (1998), the range across 52 studies was from 1% to 51%. When studies were clustered across age groups, the median rates of behavioural disorders were 8% in pre-schoolers, 12% in preadolescents, and 15% in adolescents. Results vary because of the method of sampling, how psychopathology is defined, and the cut-point used to indicate psychopathology.

If we are to analyze the effect of the context of schooling on students' mental health and emotional well-being, a useful conceptualization of school context comes from Eccles and Roeser (2011). They view it as a bridge between the macro-level of society and culture that shapes district policies and the practice of education from afar, and the middle-and micro-levels of the district, the school as an organization, and the classrooms within a school whose people, through daily acts of leadership, teaching, and social interaction, affects children's and adolescents' learning and development in immediate ways (Eccles and

Roeser, 2011, p. 255; Roeser et al., 2009). In this study, we focus on the on the middle- and micro-levels of the context of schooling, looking in particular at teaching approaches and effectiveness, teachers' expectations for success and relationships with students, and classroom disciplinary climate. Moreover, we observe at the micro-level the interaction between students' perception of the challenge of the curriculum and students' toolkit of skills and abilities to cope with it. The learning process emerges from the tension between exposure to new, and therefore challenging, material and activation of solving and decoding skills. At the same time, students' anxiety and depression stem or are exacerbated from the mismatch between challenge and skills. To formally approach the study of the micro-level interactions between students' abilities and learning challenges, we propose the use of Csikszentmihalyi's (1991) theory of flow. Csikszentmihalyi's idea of *flow* provides a useful tool for considering emotional outcomes resulting from different combinations of learning processes and students' skills. Csikszentmihalyi (1997) describes flow as deep absorption in an activity that is intrinsically interesting. Flow is believed to occur at the point of balance between the challenge inherent in the task at hand and the skills required to accomplish it. Applied to education, Csikszentmihalyi theorized four general relationships between skills and instructional challenge in students' experience of learning:

- Low-Skills/Low-Challenge – students are more likely to feel apathetic about learning because they find themselves in learning situations where they have low skills and the tasks they are asked to perform are of low-challenge. These students tend to give up because school work is inconsequential.
- Low-Skills/High-Challenge – students are more likely to feel worried in learning situations because they have low confidence in their skills and the tasks they are asked to perform are perceived as too challenging.
- High-Skills/Low-Challenge – students are more likely to feel that the challenges of learning are too few in relation to their skills, and they are unable to identify how they can make the experience more challenging. This leads to boredom and disengagement because they see little relevance in what they are asked to learn.
- High-Skills/High-Challenge – students generally feel that their skills and the challenges of the tasks they are asked to perform are in balance. These students frequently experience *flow*.

According to this theory, when students experience flow, the relationship between skills and challenge is symbiotic, where skills are neither too low nor too high in relation to the challenge at hand. Student engagement is conceived as the culmination of concentration, interest and enjoyment, as opposed to boredom or apathy (Shernoff et al., 2003).

In this study we offer estimates of the prevalence of anxiety and depression for boys and girls in middle and high schools in Canada, based on data collected from students in middle and high school using the Tell Them From Me (TTFM) survey. Students self-rated their feelings of depression and anxiety on a 3- point rating scale. We also estimate the extent to which the prevalence of anxiety and depression varies among middle and high schools, and ask whether experiencing anxiety and depression is related to children's socioeconomic status and their perceptions of skills and challenge as described by Csikszentmihalyi. Finally, we consider whether anxiety and depression are related to aspects of classroom and school climate that are known to be 'drivers' of student achievement.

We seek to answer the following questions:

1. What is the prevalence of children suffering mild to severe anxiety and depression?
2. Does the prevalence of anxiety and depression vary with grade level and between the sexes?
3. What is the relationship between anxiety and depression and students' grades?
4. What aspects of the school environment, such as teacher-student relations, the quality of instruction, classroom learning climate, expectations for success, and advocacy at school, are related to levels of anxiety and depression, after controlling for students' family socioeconomic status?

2. Methods

Data from the Fall term of 2011 Tell Them From Me (TTFM) survey were analyzed applying hierarchical linear modelling (HLM) for dichotomous outcomes (Raudenbush & Byrk, 2002). The original dataset consisted of 277,836 children attending 802 schools, and enrolled in grades 4 to 12. For the purpose of this study, we focused on 272,987 students attending grades 6 to 8 (middle school) and grades 9 to 12 (high school). Multilevel analyses were weighted at level 1 in order for students to match the socioeconomic distribution of the Canadian student population.

The TTFM measures for anxiety and depression are based on sets of six Likert items each. For example, the items for anxiety pertain to children's feelings of fear, intense anxiety, and worry about particular social events or situations, while the items for depression pertain to children's feelings of sadness, discouragement and inadequacy, and a failure to experience joy and happiness in activities at school and at home. For each item, students had to assess the frequency of that experience on a 4-point scale: 0 = "Never or hardly ever", 1 = "About once a week", 2 = "About 2 or 3 times a week", and 3 = "Every day or almost every day". Therefore the raw scores range from 0 to 18 on each scale.

The raw scores were transformed into continuous variables for anxiety and depression using a Samejima Item Response Theory (IRT) model (Hambleton et al., 1991; Linden & Hambleton, 1997). For the continuous measure of anxiety, a cut-off score was set to classify students as experiencing anxiety *versus* not experiencing anxiety. The cut-off score corresponds to an IRT score for a student that responded "Every day or almost every day" to the three easiest items and answered "About 2 or 3 times a week" to the three most difficult items. The same approach was used for classifying students as depressed *versus* not depressed.

A typology based on Csikszentmihalyi's four general relationships between skills and instructional challenge in students' experience of learning was created by means of a 2x2 tabulation. The instructional challenge dimension was assessed via six Likert items expressing the extent to which students felt challenged in their language arts, mathematics, and science classes. The scores were averaged across the three subjects and scaled on a ten-point scale. Students with scores at or above 5.0 (i.e., neutral or higher) were considered to have a high level of challenge, while those with scores below 5.0 were considered "low challenge". The skill dimension was assessed by students' GPA in language arts, mathematics, and science classes. Students obtaining an average GPA of 3.0 and above were labeled as having high skills, while those with scores below 3.0 were considered to have low skills.

The analysis also used the TTFM measure of socioeconomic status which is derived from student data on parents' education, a set of educational and cultural possessions in the home, and whether the student was living in a one-parent or two-parent family.

Several studies have shown that the learning climate of the school has an effect on students' academic achievement, even after taking account of students' family backgrounds (Willms, 2010). From its inception, TTFM survey has included a number of aspects of school climate that "drive" schooling outcomes, based on findings from a number of large-scale national and international assessments. These include quality of instruction, teacher/student relations, the disciplinary climate of the classroom, and expectations for academic success (for comprehensive reviews, see: Rutter, 1983; Sammons, Hillman, & Mortimore, 1995; Scheerens, 1992). The survey also includes a measure of student advocacy at school. However, very few studies have examined whether students' anxiety and depression are also related to these key aspects of classroom and school climate. Our analyses discern whether these factors are related to student anxiety and depression after controlling for students' family background.

3. Results

Table 1 reports the numbers of schools and students enrolled in each grade. On average the sample includes about 45,000 children per grade, with the exception of grade 6, which includes nearly 10,000 children.

For each grade, still with the exception of grade 6, students are nested in about 600 schools.

Grades	Number of students	Number of schools
6	9962	416
7	39455	667
8	40016	691
9	50270	649
10	50573	573
11	45660	534
12	37051	516

Table 1. Description of sample by grades

Hierarchical linear models were fit separately to data for the sub-samples of students attending middle and high schools. These included 41,715 students nested in 432 middle schools, and 72,818 students nested in 446 high schools.

The prevalence of students suffering from anxiety and depression changes slightly by grade. Table 2 shows an increase in anxiety around grade 9 and 10, when children transition from middle to high school. The trend of depression is different: results show that there is a small but steady increase in reporting depressive feelings from grade 6 to grade 12.

The average prevalence across grades is about 7% for anxiety and 8% for depression

Grade	Anxiety	Depression
6	6.56	6.32
7	7.00	6.49
8	7.18	7.62
9	7.08	8.08
10	6.83	8.68
11	6.48	8.61
12	6.39	8.77
Total	6.80	8.03

Table 2. Prevalence of children suffering from anxiety and depression by grade

In Figure 1, we compare the prevalence of emotionally vulnerable students between middle and high school. With separate multilevel logistic regression models, we estimate the percentage of students classified as anxious and depressed in each school.

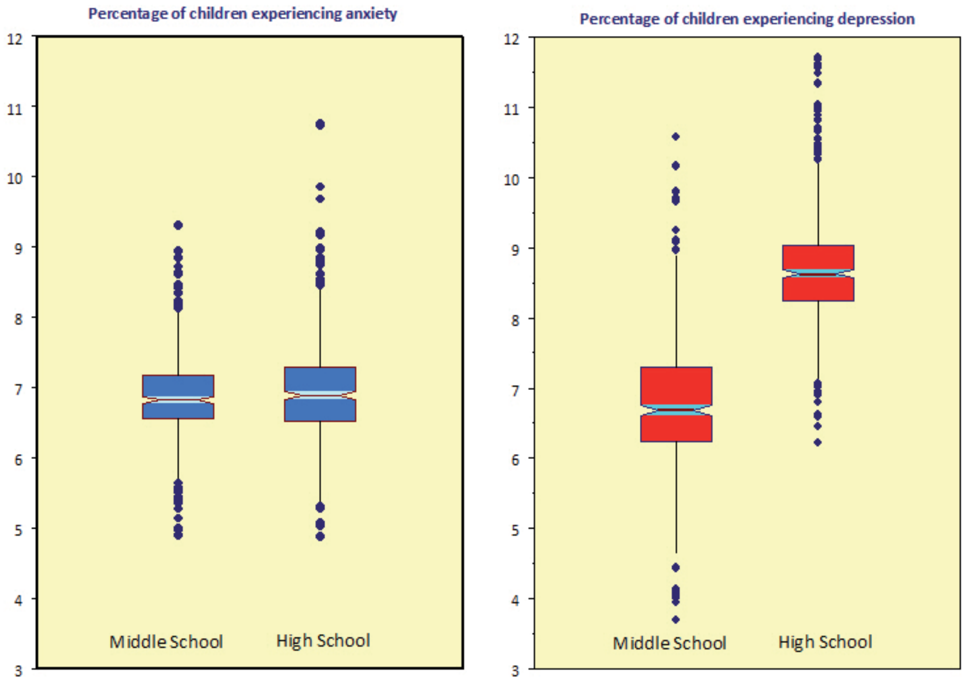


Fig. 1. Prevalence of vulnerable students in middle and high school

The graphs show that there is a significant difference in the median prevalence of vulnerable students in middle and high schools. The graphs also show the interquartile range in the prevalence of anxiety and depression for each type of school, and the presence of outliers.

Schools vary substantially in the prevalence of both anxiety and depression. While the median prevalence of anxiety is similar for middle and high schools, the range is greater for high schools. For depression the median prevalence is considerably higher for high schools than for middle schools – 6.7% compared to 8.9%. However, there is greater variability in the prevalence among middle schools.

Our next question is whether the prevalence of anxiety and depression varies as students make their way through middle and high school. Figures 2 and 3 show the trends in anxiety and depression by grade and sex.

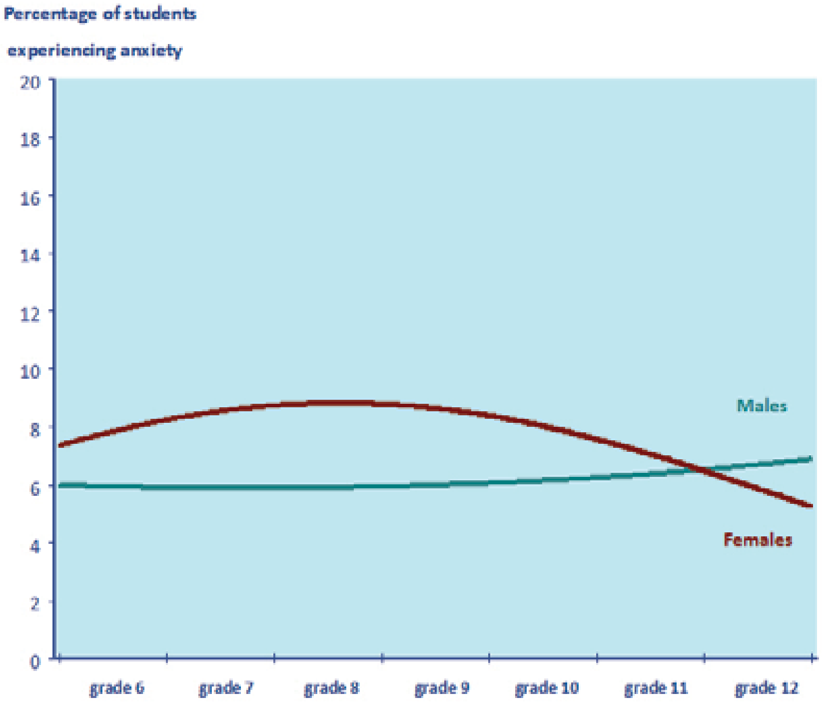


Fig. 2. Prevalence of anxiety by grade and sex

Throughout middle and high school, girls report consistently higher levels of anxiety than boys, particularly those enrolled in the last grade of middle school and those who have just entered high school. However, the prevalence of anxious girls declines by the end of grade 11, descending below the level of anxiety for boys.

The trend for depression is similar to that of anxiety, with the prevalence for girls higher than that of boys. However, the prevalence for both sexes increases considerably throughout the middle and high school years, and after grade 10, the prevalence for girls falls through to the end of high school.

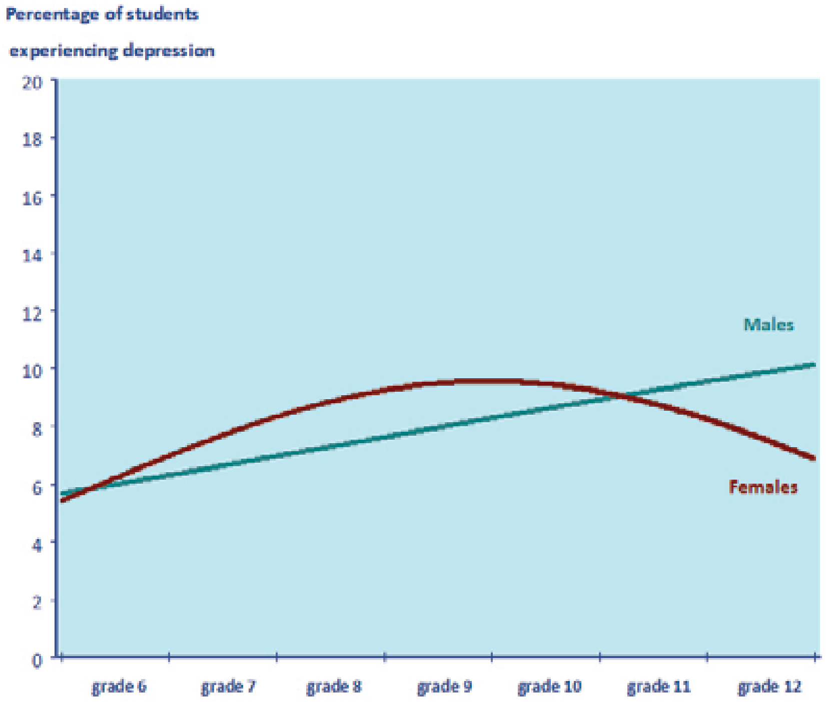


Fig. 3. Prevalence of depression by grade and sex

The next set of analyses examines the relationships between students' anxiety and depression and their learning experience in school. In Table 3, we report the percentage of students in each of the four quadrants of Csikszentmihalyi's typology, distinguishing between sex and type of school.

In middle schools, about 44% of girls and 45% of boys are in the desirable "flow" quadrant, with high challenge and high skills. For girls, 23% have low skills, and of these about two-thirds experience high levels of challenge while one-third experience low challenge. About one-third of all girls reported high levels of skills with low levels of challenge. For boys, a larger percentage – 31% – have low skills. About three-quarters of these boys reported high levels of challenge, with the remaining one-quarter – about 8% of the full middle school sample – reporting low skills and low challenge. About one-quarter of all middle school boys were in the high skills, low challenge quadrant.

The pattern for high school students is similar, but with a higher percentage of girls with low skills. About 22% of girls and 34% of boys were in the high skills, low challenge quadrant.

	Percentage of Students			
	Females		**Males**	
	Low Grades	**High Grades**	**Low Grades**	**High Grades**
Middle Schools				
High Challenge	16.1	43.9	23.4	45.0
Low Challenge	6.9	33.2	7.8	23.8
High Schools				
High Challenge	27.0	38.9	17.0	35.4
Low Challenge	11.8	22.3	13.2	34.4

Table 3. Percentage of students in the four quadrants of Csikszentmihalyi's typology, by school type and students' sex

Tables 4 and 5 report the Odds-Ratios for experiencing anxiety and depression associated with sex and membership in each of the four quadrants of Csikszentmihalyi's typology. Boys *in flow* are the reference category.

	Middle school	**High school**
Grades-Challenge Profile		
High Grades-High Challenge (reference) (HG-HC)	**1.00**	**1.00**
High Grades-Low Challenge (HG-LC)	0.90	**0.82**
Low Grades-High Challenge (LG-HC)	**1.79**	**1.30**
Low Grades-Low Challenge (LG-LC)	**1.56**	**1.38**
Female	**1.30**	**1.24**
Female * HG-LC	1.00	**1.27**
Female * LG-HC	**1.32**	**1.36**
Female * LG-LC	**1.33**	**1.23**

Note. Results in bold text are statistically significant ($p < 0.05$).

Table 4. Odds-Ratios for anxiety associated with the grades-challenge profile and sex

In middle schools, the odds of experiencing anxiety for a boy in the high grades, low challenge quadrant do not differ from a boy in flow – the odds ratio is 0.90 and is not statistically significant. However, boys in the two low skill quadrants are significantly more likely to experience anxiety, especially those in the low grades - high challenge group; their odds are 1.79 times that of boys in flow. Girls in flow are more likely to experience anxiety than boys in flow – the odds ratio is 1.30. Also, the increased anxiety associated with low skills is even more pronounced for girls than boys. For example, the odds of experiencing anxiety for girls with low grades, high challenge are (1.79 * 1.32 =) 2.36 times that of boys in flow, while girls with low grades, low challenge had increased odds of (1.56 * 1.33 =) 1.92. Girls in the high grades, low challenge group have the same likelihood of experiencing anxiety as boys in flow.

The same general pattern is evident at the high school level, although the effects are not as pronounced. Girls in flow are more likely to experience anxiety than boys in flow, and boys in the two low skills quadrants are also at increased risk of experiencing anxiety. For girls with low skills, the effect is even greater.

	Middle school	**High school**
Grades-Challenge Profile		
High Grades-High Challenge (reference) (HG-HC)	**1.00**	**1.00**
High Grades-Low Challenge (HG-LC)	0.91	1.04
Low Grades-High Challenge (LG-HC)	**1.60**	**1.31**
Low Grades-Low Challenge (LG-LC)	**1.77**	**1.65**
Female	0.97	1.01
Female * HG-LC	1.11	**1.22**
Female * LG-HC	**1.59**	**1.36**
Female * LG-LC	**1.36**	**1.20**

Note. Results in bold text are statistically significant ($p < 0.05$).

Table 5. Odds-Ratios for depression associated with the grades-challenge profile and gender

Table 5 shows the results for depression. Middle school boys and girls who have low grades, irrespective of the challenge they have to deal with, are more prone to experiencing depression than those with high grades. For boys the odds are greater by a factor of 1.60 for those with high challenge and 1.77 for those with low challenge. The effect associated with low skills is even more pronounced for girls, with odds ratios of (1.60 * 1.59 =) 2.54 and (1.77 * 1.36 =) 2.41 for high and low challenge respectively.

Similar effects are observed at the high school level, although the effects are not as strong. Students with low skills are more likely to experience depression. This effect is stronger for boys than for girls.

	Anxiety		**Depression**	
	Middle school	**High school**	**Middle school**	**High school**
Student-level				
Socioeconomic status	**0.79**	**0.78**	**0.76**	**0.79**
Expectations for success	**0.87**	**0.86**	**0.80**	**0.82**
Advocacy at school	**1.03**	**1.03**	**1.02**	1.01
School-level				
Quality of teaching	1.14	1.21	1.04	**1.25**
Classroom disciplinary climate	0.92	1.07	0.88	1.06
Teacher-student relations	0.86	**0.80**	0.95	**0.82**

Note. Results in bold text are statistically significant ($p < 0.05$).

Table 6. Odds-Ratios for experiencing anxiety and depression associated with school processes and socioeconomic status

The last table examines the relationship between anxiety and depression with socioeconomic status and the measures of classroom and school context. The results were derived from a multilevel analysis with SES at the child level and the classroom and school context factors measured as the school level. It was not possible to identify students' classroom membership with the TTFM data. Students with higher SES are less prone to experiencing anxiety and depression at both the middle and high school levels. The odds-ratios are very close, approximately 0.75, for both anxiety and depression at both levels of schooling. This indicates that a student with an SES that is 0.5 standard deviations above the mean (i.e., about the 69th percentile) is only three-quarters as likely to experience anxiety or depression as a student with an SES that is 0.5 standard deviations below the mean (i.e., at the 31st percentile).

Advocacy at school and expectations for success were also entered as student-level variables in the model. The argument is that it is students' own experience of teachers' expectations of them and their own sense that they have an advocate at school that is important to their well-being. Teachers' expectations for success appear as a protective factor. In middle schools, each one-point increase on the ten-point scale is associated with a decrease in the odds of a student experiencing anxiety or depression. This relationship is stronger in high schools than in middle schools. A student's report of having an advocate at school has a very weak relationship with anxiety and depression; the odds-ratios are very close to one at both levels of schooling.

Two school process variables, classroom disciplinary climate and teacher-student relations, which were measured at the school level, also appear as protective factors. A one-point increase on the ten-point classroom disciplinary climate scale is associated with a decrease in the odds of anxiety of about 14% at the middle school level and by about 6% at the high school level. However, these results were not statistically significant. Slightly stronger effects were observed for depression, but in this case also the results were not statistically significant. Similar findings were evident for teacher-student relations, although in this case the protective effect was significant for both anxiety and depression at the secondary level. The odds-ratios for quality of teaching were greater than 1.0 for anxiety, indicating that it is a risk factor for anxiety and depression. However, these estimates were not statistically significant. For depression, the odds-ratio was 1.22 and statistically significant at the high school level. This counter-veiling effect is comparable to the protective effect of teacher-student relations.

4. Discussion

This paper examined the prevalence of anxiety and depression between males and females for a large sample of students that completed the *Tell Them From Me* student survey in the fall of 2011. *Tell Them From Me* is Canada's largest school survey, covering a wide range of student outcomes as well as school and classroom processes. Four key findings emerged from the analysis.

First, girls have a higher prevalence of anxiety and depression than boys in middle school and in the early stages of high school. The gap is largest at the time when students make the transition from middle to high school. The gender gap at the end of grade 8 is slightly smaller for depression than it is for anxiety. Another key finding is that the prevalence of depression does not decline during the high school years for boys as it does for girls; rather it increases linearly through to the end of grade 12, when it reaches about 10%.

The results of this study show clearly why the prevalence of depression, in absolute terms, can vary substantially across studies. One source of variability is sampling error, which is largely ruled out in this study due to the large sample size. However, many studies reported in the literature are based on samples with less than 1000 students. Another source is the age range considered. These findings show clearly that the results can range considerably depending on the grade or age level considered; this is especially the case for depression. A third source of variation is measurement error. This is affected not only by the nature and quality of the items in the scales but also by the cut-point on the scale used to denote psychopathology. We developed separate continuous measures for each scale, anxiety and depression, using a variant of Item Response Theory for graded responses (Samejima, 1997). The approach takes into account the 'difficulty' of each item as well as the response categories for each item. For example, an anxiety item about whether students are concerned about what other students think of them is 'easier' than an item about feeling fearful and nervous, and thus has a lower value on the IRT scale. Also, each item has four separate IRT scores, corresponding to the frequency reported by a student (e.g., "Never or hardly ever", "About once a week", "About 2 or 3 times a week", or "Every day or almost every day"). Therefore, each pattern of scores yields a different score on the IRT scale. We set the cut-point at an IRT corresponding to a pattern in which a student answered "Every day or almost every day" to the three easiest items and answered "About 2 or 3 times a week" to the three most difficult items.

Thus, the absolute prevalence is arbitrary, as we would have reported a different prevalence if we had set the cut-point corresponding to a different pattern, such as a student that answered "Every day or almost every day" to all six items. The prevalence would also change if we added an easier or more difficult item to the scale. Quite often, researchers hold up clinical diagnosis as the 'gold standard'. However, clinicians also have their own internal 'cut-points' for discerning whether or not a child is anxious or depressed, and there is considerable variation among clinicians (Roberts et al., 1998). In school-based studies in which the prevalence of mental health issues is derived from reported cases, the prevalence can also vary because of the availability of clinicians in certain geographic areas.

Thus, although the estimate of prevalence depends heavily on the cut-point set to define psychopathology, our scaled-based approach has several advantages. First, the meaning of the threshold is anchored to the response categories of the items. Second, it can be applied to any scale, regardless of whether there are three, four, or more response categories, and any number of items. Third, the approach can be applied across all children in the sample, regardless of age, sex, or other characteristics. One can also assess whether the set of items is 'measurement invariant' for age, sex, or other factors (Wu, Li, and Zumbo, 2007). An important aspect of measurement invariance is whether there are certain items that are 'easier' for girls than boys, or vice-versa, given their overall level of anxiety or depression. Fourth, which we see an a key advantage for this type of study, it provides a reliable method for making comparisons among sub-populations, across jurisdictions such as schools and school districts, and over time. For practical purposes, these are the kinds of comparisons that are relevant to schools and school districts that use TTFM.

The second key finding is that most middle and high schools have a prevalence of anxiety that is close to the national median. The inter-quartile range is only about 1% for both anxiety and depression. However, there are many notable outliers, with decidedly low or high levels of anxiety and depression. These are not simply due to measurement or sampling error, as TTFM assesses all children within schools and our HLM analysis provides "shrunken" estimates of each school's prevalence which takes into account measurement and sampling error. For educators, this means that most interventions need to focus on within-school strategies to reduce anxiety and depression, but there are some schools that warrant a whole-school, targeted approach. For researchers, these results raise questions about why some schools have markedly low levels of anxiety and depression while other have very high levels.

Third, about 35 to 45% of Canadian students were considered to be in "flow"; that is, have strong skills and feel they were challenged in their classes. This general finding was reported earlier by the Canadian Education Association as part of its project called, "What did you do in school today?"(Willms et al., 2009), which has spawned national interest in the extent to which children are intellectually engaged at school. This study shows that these students are less prone to experiencing anxiety or depression. However, students with high skills but lack challenge also have relatively low levels of anxiety and depression. While these students may lack intellectual engagement, it does not appear to increase anxiety or cause depression. Rather, our findings suggest that skills, not challenge, is the critical aspect in the Csikszentmihalyi schema. Moreover, there is a significant gender by skill interaction:

girls with low levels of skills tend to be more at risk than boys experiencing anxiety and depression.

An earlier study based on TTFM data used students' perceptions of their skills as the skill measure in the Csikszentmihalyi schema (Tramonte & Willms, 2010), while this study used a measure of skill derived from students' reports of grades. Analyses by The Learning Bar Inc. have shown that grades are a more reliable measure of skills than students' perceptions, but the findings from the earlier Tramonte and Willms (2010) are similar to those found in this study. As a factor that contributes to anxiety and depression, it seems that students' perceptions of their skills in relation to the peers in their context are equally important to students' achieved grades.

Our fourth key finding is also concerned with why schools vary in their levels of anxiety and depression. We approached this using a traditional multilevel analysis of school effects, using measures of schooling processes that have been found in many studies to be related to students' academic achievement. This strategy was only moderately successful. Students' SES and teachers' expectations for success proved to be strong student-level protective factors; however, students having an advocate at school had weak relationships with anxiety and depression. At the school level, classroom disciplinary climate and teacher-student relations had protective effects, as expected, but the analysis lacked statistical power and thus some of the estimates were not statistically significant. In hierarchical analyses, the statistical power of school-level factors depends more on the number of schools than the number of students in this type of study. The effects of quality of teaching were in the opposite direction as expected. This likely stems from collinearity at the school level in these factors; when quality of instruction is entered separately, without classroom disciplinary climate and teacher-student relations in the model, the results are weak but in the expected direction. As with many of the traditional 'school effects' analyses aimed at uncovering the effects of school processes on achievement, it is difficult to identify the key processes that make a difference. The results call for controlled intervention studies with longitudinal designs.

5. Conclusion

The findings of this study provide a compelling argument that the mental health of adolescents deserves equal or even greater attention than academic achievement in discussions about school effectiveness. Kessler et al. (2005) report that about one-half of all Americans will experience a mental health condition at some time during their lives, and in most cases the first onset is during childhood or adolescence. The *Tell Them From Me* evaluation system provides reliable data on the prevalence of anxiety and depression, enabling schools to track their progress on key outcomes. Students with low levels of skills are more prone to anxiety and depression, a relationship which is evident for most students when they begin middle school. Schools can make a difference, not only through staff efforts at helping students at the individual level, but also by establishing a positive school climate characterized by high expectations for success, a positive classroom disciplinary climate, and strong teacher-student relationships.

6. References

Bagnell, A., Tramonte, L. & Willms, J. D. (2008). *The Prevalence of Significant Mental Health Problems among Canadian Youth and their Comorbidity with Cognitive and Health Problems*. Human Resources and Skills Development Canada: Ottawa, ON.

Blanchard, L. T., Gurka, M. J. & Blackman, J. A. (2006) Emotional, developmental, and behavioral health of American children and their families: A report from the 2003 National Survey of Children's Health. *Paediatrics*, Vol. 117, No 6, (June, 2006), p.p. (1202-1212), 0031-400

Bowlby, J.W. & McMullen, K. (2003) At a Crossroads: First Results for the 18 to 20-Year-old Cohort of the Youth in Transition Survey. Human Resources Development Canada, 0-662-31555-3, Ottawa: ON

Cole, T. J., Bellizzi, M. c., Fegal, K. M. & Dietz, W. H. (2000) Establishing a standard definition of overweight and obesity worldwide: International survey. *BMJ,* Vol. 320, (May, 2000), pp. (1240-1243), 10.1136

Csikszentmihalyi, M. (1991). *Flow: The Psychology of Optimal Experience*. Harper & Row Publishers, 978-0060920432, New York.

Csikszentmihalyi, M. (1997). *Finding Flow: The Psychology of Engagement with Everyday Life*. Harper & Row Publishers, 978-0465024117, New York.

Eccles, J. S. & Roeser, R. W. (2010). An ecological view of schools and development. In *Handbook of Research on Schools, Schooling and Human Development,* J. L Meece and J. S. Eccles, pp. (6-21). Routledge, 978-0805859492, New York .

Eccles, J. S. and Roeser, R. W. (2011), Schools as Developmental Contexts During Adolescence. *Journal of Research on Adolescence*, Vol. 21, No. 1, (March, 2011), p.p. (225–241), 10508392

Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991) *Fundamentals of Item Response Theory.* Sage Press, 978-0803936478, Newbury Park, CA.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R. & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry*; Vol. 62, No. 6, (June, 2005), pp. (593-602), 0003-990X

Kolko, D. J. & Kazdin, A. E. (1993). "Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent, and teacher reports." *Journal of Child Psychology & Allied Disciples*, Vol. 34, No. 6, pp. (991-1006), 1469-7610

Linden, W. V. & Hambleton, R. K. (1997). *Handbook of Modern Item Response Theory*. Springer, 978-0387946610, New York.

Murray, C. J., & Lopez, A. D. (Eds.). (1996). The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Harvard School of Public Health on behalf of the World Health Organization. Harvard University Press, Cambridge, MA.

Pope, D. C. (2001). *"Doing school": How we are Creating a Generation of Stressed out, Materialistic, and Miseducated Students*. Yale University Press, 978-0300098334, New Haven.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models (2 ed.)*. Sage Publications Inc., 978-0761919049, Newbury Park, CA.

Roberts, R. E., Attkisson, C. C. & Rosenblatt, A. (1998). Prevalence of psychopathology among children and adolescents. *American Journal of Psychiatry*, Vol. 155, No. 6, (June, 1998), pp.(715–725), 9619142

Roeser, R.W., Urdan, T. C. & Stephens, J.M. (2009). School as a context of motivation and development. In *Handbook of Motivation at School,* K.R. Wentzel & A. Wigfield, pp. (381-410). Routledge, 978-0805862843, New York.

Rutter, M. (1983). School Effects on Pupil Progress: Research findings and policy implications. *Child Development*, Vol. 54, pp. (1-29).

Samejima, F. (1997). Graded response model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory* (pp. 85-100). New York: Springer.

Sammons, P., Hillman, L. and Mortimore, P. (1995) Key Characteristics of Effective Schools: A Review of School Effectiveness Research. London: Institute of Education.

Sanford, M. N., Offord, D. R., Boyle, M. H., Peace, A., & Racine, Y. A. (1992). Ontario Child Health Study: Social and school impairments in children aged 6 to 16 years. *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 31, No. 1, (January, 1992), pp. (60-67), 0890-8567

Seligman, M.E.P., Walker, E., & Rosenhan, D.L. (2001). *Abnormal Psychology,* (4th ed.) W.W. Norton, 978-0393944594, New York.

Scheerens, J. (1992). *Effective Schooling, Research, Theory and Practice*. London: Cassell.

Shernoff, D. J., Csikszentmiahlyi, M., Schneider, B. & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psych Quart*, Vol. 18, No. 2, (June, 2003), p.p. (158–176), 10453830

Sznitman, S.R., Reisel, L., & Romer, D. (2011) The Neglected Role of Adolescent Emotional Well-Being in National Educational Achievement: Bridging the Gap Between Education and Mental Health Policies. *Journal of Adolescent Health*, Vol. 48, No. 2, (Febuary, 2011), p.p, (135–142), 1054-139x

Stark, K. D. & Smith, A. (1995). Cognitive and behavioral treatment of childhood depression. In *Behavioral Approaches for Children and Adolescents: Challenges for the Next Century*, H.P.G.van Bilsen, P.C.Kendall and J.H.Slavenburg, pp. (113-144), Plenum Press, 9780306451225, New York.

Tramonte, L., & Willms, J.D. (2010). The prevalence of anxiety among middle and secondary school students in Canada. *Canadian Journal of Public Health*, Vol. 101, No. 3, pp. (19-23), 00084263.

Waddell, C., Offord, D. R., Shepherd, C. A., Hua, J. M., & McEwan, K. (2002). Child psychiatric epidemiology and Canadian public policy-making: The state of the science and the art of the possible. *Canadian Journal of Psychiatry,* Vol. 47, No. 9, (November, 2002), pp. 825-832, 0706-7437

Wang, P. S., Berglund, P., Olfson, M., Pincus, H. A., Wells, K. B. & Kessler, R. C. (2005) Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry*, Vol. 62, No. 60, (June, 2005), p.p. (603-613), 0003-990X

Willms, J. D. (2003) *Student Engagement at School: a Sense of Belonging and Participation: Results from PISA 2000*. Organization for Economic Co-operation and Development, Paris.

Willlms, J. D. (2006). Learning Divides: Ten Policy Questions about the Performance and Equity of Schools and Schooling Systems. UNESCO Institute for Statistics, 92-9189-039-1, Montreal, QC.

Willms, J. D. (2008). *Developmental Trajectories of Canadian Children and Youth*. Human Resources and Skills Development Canada, Ottawa, ON.

Willms, J.D. (2010). *School* composition and contextual effects on student outcomes. *Teachers College Record. Vol.* 112, No 4, pp. (1,008-37).

Willms, J. D., Friesen, F. & Milton, P. (2009). *What Did you Do in School Today? Transforming Classrooms Through Social, Academic, and Intellectual Engagement*. Canadian Education Association, 1-896660-36-3, Toronto: ON.

Willms, J. D., Tramonte, L., & Chin, N. (2007). *The Use of Item Response Theory for Scaling Behavioural Outcomes in the NLSCY for the Successful Transitions Project.* Unpublished report. Ottawa: Human Resources and Skills Development Canada

Wu, A. D., Li, Z., & Zumbo B. D., (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS Data. *Practical Assessment Research & Evaluation, 12 (3),* 1-26.

23

Re-Emergence of HIV Infection and Syphilis Among Men Who Have Sex with Men

Maria Antonella Di Benedetto, Nino Romano and Alberto Firenze
Dipartimento di Scienze per la Promozione della Salute, University of Palermo
Italy

1. Introduction

Sexual activity has been shown to be the primary mode of transmission for several important viral and bacterial infections among men who have sex with men (MSM) throughout the world and therefore sexually transmitted infections (STIs) are still one of major public health problem worldwide. Since the earliest days of the human immunodeficiency virus (HIV) epidemic MSM have been profoundly affected by the disease; by the time the first AIDS cases were reported in the early 1980s, HIV had already spread widely among homosexual and drug-user communities. In western Europe, has been estimated that HIV incidence peaked around 1983 among MSM and in 1987-88 among injecting drug users, with 120.000 homosexual/bisexual men infected by 1985, and 144.000 injecting drug users infected by 1989 (Downs et al., 2000). MSM account for a substantial proportion of HIV infections and compose a "bridging group" for transmission to heterosexuals because of the high frequency of reported bisexuality. Widescale behavioural modification in response to the emerging HIV/AIDS pandemic saw dramatic reductions in the incidence of many acute STIs in the late 1980s (WHO, 2007). However, these have not been maintained and many states are now observing increases in rates of diagnosed STIs; then the proportion of homosexual men reporting "unsafe sex", often measured as unprotected anal intercourse (UAI) with casual partners, has increased since the mid-1990s (Prestage et al., 2005). In western Europe, among MSM, HIV diagnoses increased in 2002 (22%, from 2757 in 2001 to 3371 in 2002) after a slow decline in the previous years. Increased high risk sexual behaviour and migration of HIV-infected people from sub-Saharan Africa have contributed to the rise in the incidence of STIs since 2000 (Hamers & Downs, 2004). On the other hand, antiretroviral treatment, which lowers the viral load, may reduce the likelihood that an HIV infected person infect a partner during unprotected sexual intercourse. Mathematical modelling suggests that the effect of antiretroviral treatment on lowering transmission at the population level could be offset by an increase, even modest, in risk behaviour, and the net effect of these opposing forces on HIV incidence has been much debated (Katz et al., 2002).

This HIV infection increase, in some industrialized countries, has also coincided with several reported outbreaks of syphilis among the MSM population (Couturier et al., 2004); in particular, oral sex between men appears to have been important in the transmission of syphilis in outbreaks reported overseas, probably because among MSM unprotected oral sex is considered a safe practice. Data of literature report that syphilis transmission attributed to oral sex has been estimated between 20 and 46% (Peterman & Furness, 2007). Syphilis remains a global problem despite the existence of simple and validated screening tests; moreover, it has been described as strongly associated with prevalent HIV infection among MSM thereby playing an important role in the increase of this infection in a susceptible population. The interaction between HIV infection and syphilis is of great interest for the epidemiological and clinical implications that apply to both heterosexuals and MSM (Peterman & Furness, 2007). The World Health Organization (WHO) estimates 12 million new syphilis cases worldwide annually, of which 140.000 in western Europe where the infection has re-emerged both in homosexual men, in particular in HIV-positive gay men, and in heterosexuals (WHO, 2001; Simms et al., 2005; Xu et al., 2011). An increase of cases among MSM has been found in Italy (Istituto Nazionale di Statistica [ISTAT], 2005), Sweden (Velicko et al., 2004), Scotland (Wallace et al., 2005) and in Germany (Marcus & Hamouda, 2004), suggesting that syphilis is now endemic in the MSM population in parts of Western Europe. In the USA, the scenario is similar because an increase in syphilis cases among MSM occurred between 2000 and 2005, characterised by high rates of HIV co-infection and high-risk sexual behaviour (Fenton et al., 2008). As infection is frequently asymptomatic, annual serological screening for syphilis is to recommend in this group of men. Syphilis offers a classic example of the re-emergence of an easily diagnosed and readily treatable STD.

In this report we review the recent epidemiology of HIV infection and syphilis among MSM in the developed and developing world and discuss potential reasons for the emerging trends and their implications for prevention and control.

2. HIV /AIDS

HIV infects defense/immune system cells such as CD4+ T cells, macrophages and dendritic cells. After infection HIV uses CD4+ cells as host to make copies and infect other cells, this leads to the reduction of CD4+ cells and immune system totally collapse. The development from HIV to AIDS is checked by decline of CD4+ cells. Two types of HIV has been characterized, HIV-1 and HIV-2. HIV-1 is the most virulent and pathogenic strain; the HIV-2 epidemic remains limited, both in terms of prevalence and with respect to geographic spread because, worldwide, the predominant virus is HIV-1. The relatively uncommon HIV-2 type is concentrated in West Africa (Senegal, Ivory Coast, Guinea Bissau) and other countries characterized by a considerable immigration of people coming from West Africa. The HIV-1 is further divided into 3 groups: 1. group M (main) ; 2. group O (outlier) ; 3. group N (non M - non O). These groups have been identified in there envelop region. The major HIV pandemic is caused by HIV-1 group M strains; groups O and N of HIV-1 are rare and limited to Cameroon in Central Africa (McCutchan, 2006). Group O accounts for less than 10% of HIV-1 infections in Cameroon and has spread little beyond these confines ; only a handful of group N infections have been identified. Within HIV-1 group M, nine subtypes are recognized namely subtypes A, B, C, D, F, G, H, J and K, that represent lineages of HIV and have different geographic location. Many of the HIV-1 group M subtypes are rare and remain at low prevalence. In contrast, subtype B is widespread globally, dominating epidemics in the Americas, Western Europe, and Australia. Subtype C accounts for more than 50% of all infections worldwide, concentrated in Southern and East Africa, and in India; subtype D strains are found principally in East Africa, and to a lesser extent in West Africa. In some parts of the world less information is available about HIV diversity, particularly North Africa, the Middle East, and parts of Central Asia (McCutchan, 2006).

The study and monitoring of the genetic evolution of the HIV-1 represent an essential strategy for controlling the global HIV-1 epidemic and for developing efficient preventive and therapeutic strategies.

2.1 The current scenario in developed countries

The HIV epidemic is resurging in gay communities in developed world. As a consequence, a number of European countries have conducted behavioural surveys among MSM to monitor HIV and STIs risk in this population ; behavioural surveillance provides important information for planning and evaluating prevention interventions (Elford et al., 2009). In Canada (Public Health Agency of Canada, 2007), the United States (CDC, 2007) and Australia (National Centre in HIV Epidemiology and Clinical Research, 2007) MSM represent the risk group diagnosed most frequently with HIV infection. In particular, in the United States, the MSM population accounted for the majority of new HIV infections in 2006. This re-emergence has led some commentators to suggest that HIV prevention in MSM is faltering so president Obama's released National HIV/AIDS Strategy calls for more emphasis on addressing the HIV prevention and care needs of MSM (The White House Office of National AIDS Policy, 2010).

2.1.1 European Union and European Free Trade Association countries

The number of newly diagnosed HIV cases reported among men who have sex with men (MSM) has recently increased throughout European Union (EU) and European Free Trade Association (EFTA) countries. The predominant mode of transmission for HIV infection in EU and EFTA countries is sex between men. In 2006, 7693 newly diagnosed HIV infections among MSM were reported (56.7 per million men aged 15-64 years). In most of the European countries with data for 2000-2006, the number of new HIV diagnoses increased by 86%, the median age at HIV diagnosis remained unchanged (36 years), whereas the proportion of MSM presenting with an AIDS-defining illness at the time of HIV diagnosis declined from 25% in 2000 to 10% in 2006. Basically there is a decrease of cases of AIDS and death in the presence of an increase of HIV infection. (Likatavicius et al., 2008). In general, reported HIV prevalence ranged between 8% and 68% among MSM with sexually transmitted infections, between 10% and 18% among those recruited in community settings, but remained <10% in central Europe and Ireland. Whereas the decreasing rates of AIDS diagnoses and AIDS deaths reflect relatively good access to therapy, the increasing numbers of new HIV diagnoses and relatively high prevalence of HIV among MSM suggest the need for Europe-wide HIV prevention among MSM. Currently, the prevalence of HIV among homosexual and bisexual men is 10–20% in most western European countries; prevalence is usually higher in large cities than elsewhere (Hamers & Downs, 2004). In the below report, there are summarized the results of serological investigations carried out in several European countries, which show the pattern of increasing HIV incidence in these last years. In Madrid, in a study among homosexual and bisexual men who repeatedly attended voluntary HIV testing centres, HIV incidence decreased from 4.7 per 100 person-years in 1988 to 1.1 in 1995, but then increased to 2.2 in 2000. More recently, incidence studies based on the serological testing algorithm for recent HIV seroconversion (STARHS) have been done among homosexual and bisexual men attending sexually transmitted infection clinics in the Netherlands and the UK (Hamers & Downs, 2004). In Amsterdam, an overall incidence of 3.0 infections per 100 person-years during 1991–2001 was noted, with a significant increase over time. The increase was evident in men older than 34 years but not in younger men. In the UK, the overall annual incidence was 2.4% (3.1% in London, 1.0% elsewhere) with no significant trends in HIV incidence during 1995–2001. However, more recent data showed an increasing trend to 3.5% in 2002 (Hamers & Downs, 2004). Belgium is currently experiencing an alarming upward trend among MSM in the number of new HIV diagnoses that, in the past decade, increased by 228% (from 101 cases diagnosed in 1999 to 332 cases in 2008). The majority of new HIV infections were diagnosed among Belgians citizens (72%), followed by other European nationalities, sub-Saharan Africa and other/unknown nationality (Sasse & Defraye, 2009).

In Slovenia, of a total of 48 newly diagnosed HIV infection cases reported for 2008 (23.5/million population), 34 cases, representing more than two thirds, were MSM. Since 1999, the annual reported rate of newly diagnosed HIV cases in MSM raised from 2.5 to 16.7 per million men 15-64 years old, an increase of more than six times. The overall increase in the number of newly diagnosed HIV cases in Slovenia during recent years has been due almost exclusively to the increase in new diagnoses among MSM (Klavs et al., 2009).

Increases in risky sexual behaviour in a population in which HIV is already highly prevalent, coupled with increases in sexually transmitted infections that enhance the risk of HIV transmission, have the potential to lead to an expansion of HIV transmission among homosexual and bisexual men (Hamers & Downs, 2004).

2.1.2 United States / Canada

On June, 5, 1981, MMWR published a report of *Pneumocystis carinii* pneumonia in five previously healthy young men in Los Angeles. These cases were later recognized as the first reported cases of AIDS in the United States; since that time, this disease has become one of the greatest public health in the world. HIV infection and AIDS have claimed the lives of more than 22 million persons worldwide, including more than 500.000 persons in USA (CDC, 2006). Prevention interventions have contributed to a steady decline (from an estimated 8.048 in 2001 to 5.962 in 2004) in new HIV/AIDS diagnoses (CDC, 2005) but despite these successes, several challenges remain. In the United States, in contrast to other risk groups for which HIV/AIDS diagnoses have decreased, cases among MSM increased 8,6% from 2001 up to 72% in 2006 (CDC, 2008a). Among MSM with new infections, 46% were white, 35% were black, and 19% were Hispanic. Among MSM aged 13-29 years, the number of new HIV infections in blacks (5.220) was 1.6 times the number in whites (3.330) and 2.3 times the number in Hispanics (2.300). Among black and Hispanic MSM, most new infections were in persons aged 13-29, whereas, among white MSM, most new infections were in persons aged 30-39 years (CDC, 2008b). In Milwaukee county has been reported a 144% increase during 2000-2008 in black homosexual men aged 15-29 years. Interestingly, an increase in syphilis diagnoses among young black MSM in Milwaukee preceded the increase in HIV diagnoses, which suggests that changes in risk behavior or sexual networks might explain the increase (CDC, 2011).

The Centers for Disease Control and Prevention have documented a dramatic increase in the rates of sexually transmitted diseases among men who have sex with men (MSM) living in major urban centres across the country over the past 10 years (CDC, 2004). The MSM currently represent the group with the largest number of new HIV infections in the USA and its incidence have been increasing consistently since the early 1990s. Based upon national surveillance at sexually transmitted disease clinics, median positivity rates for gonorrhoea and chlamydia are higher among HIV-infected MSM than among HIV-uninfected or unknown status MSM. Syphilis seroreactivity rates are up to four times greater among HIV-infected MSM than HIV-uninfected or unknown status MSM (http:www.cdc.gov/STD/stats06/msm.htm as cited in Mayer et al., 2010). This trend of higher comparative rates of incident sexually transmitted infection (STI) among HIV-infected MSM is also evident in community samples both nationally (Whittington, 2002, as cited in Mayer et al., 2010) and internationally (Dodds, 2007, & Dougan, 2007, as cited in Mayer et al., 2010). Consistent with high rates of STI among HIV-infected MSM, some studies suggest that rates of high-risk sexual behaviors (i.e. anal intercourse without a condom) may be up to three times higher among HIV-infected MSM compared with HIV-uninfected MSM (Mayer et al., 2010).

In Canada at the end of 2005, an estimated 58.000 people were living with HIV infection which represents an increase of about 16% from the 2002 estimate of 50.000. In terms of exposure category, prevalent infections in 2005 were comprised of 29.600 MSM (51%), 9.860 IDU (17%), 8.620 heterosexual/non-endemic (15%), 7.050 heterosexual/endemic (12%), 2.250 MSM-IDU (4%), and 400 attributed to other exposure (<1%). The largest absolute increase was among the MSM exposure category with 3.400 more prevalent infections since 2002 (13% increase) (Boulos et al., 2006). This trend among MSM and MSM-IDU is been associated with increases in risky sexual behavior. The causes of this increase are complex and may include decision-making based on false assumptions about a partner's HIV status, dissatisfaction and difficulties with condom use, feelings of marginalization, depression and the choice to not use condoms as a gesture of commitment to a partner. The scenario is different among Aboriginal persons because the distribution of new cases of HIV in 2005 was 53% IDU versus 10% MSM and 3% MSM-IDU; this highlights the uniqueness of the HIV epidemic among Aboriginal Canadians and underscores the complexity of Canada's HIV epidemic (Boulos et al., 2006). Moreover, an estimated 16.900 persons HIV positives (26% of total prevalent infections) were unaware of their HIV-infected status and this proportion varied from an estimated 19% of HIV-infected MSM, to 25% of HIV-infected people who inject drugs, and 35% of HIV-infected heterosexuals (Yang et al., 2010).

2.1.3 Australia

Since 2000 the annual number of new diagnoses of HIV in Australia has increased by 38%. Gay men continue to be the most affected, accounting for 64% of people who were newly diagnosed as having HIV infection in 2004-2008, and 82% of the infections that were determined to have been acquired no more than a year before they were diagnosed (de Wit et al., 2010). Recent trends in new HIV diagnoses differ across Australian jurisdictions (Figure 1), the region with the highest number of HIV diagnoses each year is the state of New South Wales (NSW) with approximately 70% of these cases attributed to homosexual transmission (de Wit et al., 2010). While the population rate of HIV diagnosis doubled in Victoria between 1999 and 2006, increased markedly in South Australia and Western Australia in that period and steadily increased in Queensland since 1999, in NSW over the past decade the rate has remained relatively stable (Guy et al., 2007, as cited in de Wit et al., 2010). Historically, NSW has had the highest HIV infection rate; the apparent stability of the current HIV epidemic in gay men in NSW may reflect testing patterns, HIV status knowledge and disclosure and the well-reasoned use of risk-reduction strategies that this enables. A mathematical model has been used in conjunction with HIV/AIDS data from the Australian National HIV/AIDS

Registry to estimate numbers and ages of Australian men who have sex with men living with HIV infection from 1980 to 2005 (Murray et al., 2009). The average age of HIV-infected Australian MSM is estimated to exceed 44 years of age by the year 2010 and has increased by 1 year of age for each two calendar years since the mid-1980s. HIV-infected MSM over 60 years of age have been increasing in number by 12% per year since 1995. A consequence of successful therapy with subsequent ageing of those infected has meant that from 2001 estimated deaths from other causes exceed AIDS deaths in Australia (Murray et al., 2009). Australian guidelines recommend annual testing for HIV and sexually transmitted infections (STIs) for all men who have sex with men (MSM) and 3-6 monthly testing for those at higher risk as defined by behavioural criteria; a recent study (Guy et al., 2010) showed that there is poor adherence to national guidelines that recommend regular re-testing of MSM for STIs, particularly among those at higher risk who require more frequent testing.

Clinical strategies are urgently needed to encourage more frequent HIV/STI testing among MSM, especially in the higher risk subgroup.

2.2 The current scenario in developing / transitional countries

MSM in developing/transitional countries have a several-fold higher odds of HIV infection compared to the general population, with an 18-fold higher risk in Asia (Baral et al., 2007, as cited in Solomon et al., 2010). MSM in Africa are a highly vulnerable population at risk for HIV infection, in spite of this, the Africa region has documented limited HIV and STDs data on MSM, largely due to stigmatization of homosexual behaviour in the population (Smith et al., 2009). After Asia and sub-Saharan Africa, the region most impacted by HIV infection is Latin America where high rates HIV prevalence have been documented among MSM with no evidence of significant decrease (Halperin et al., 2009).

2.2.1 India

There are more than 39.5 million people living with HIV worldwide, more than 2.5 million live in India where sexual transmission is the predominant mode of HIV transmission and is estimated to be the major risk for about 86% of those infected (National AIDS Control Organization, 2006, as cited in Solomon et al., 2010). The overall HIV prevalence in India is 0.36% ; the high reported HIV prevalence among some high-risk groups such as female sex workers (FSW) and men who have sex with men (above 5%), puts the country in the classification of concentrated epidemic. Protecting India's large population of FSW and MSM from infection is a critical priority and it is also an effective way to protect the remaining population from infection (UNAIDS, 2009, as cited in Gutierrez et al., 2010). A study carried out in Mumbai (formerly known as Bombay) found high rates of unprotected sex, bisexuality and HIV among urban Indian MSM (Kumta et al., 2010); another study conducted in Andhra Pradesh (AP) also found that MSM reported high rates of unprotected anal sex with other men and women. Andhra Pradesh (total population of 76.2 million) is among the six states in India with the highest HIV prevalence; about 10% of all AIDS cases in India are in AP (Dandona et al., 2005).

In India, the first evidence of HIV was described among sex workers in Tamil Nadu (TN) (Simoes et al., 1987, as cited in Solomon et al., 2010), then spreading to monogamous housewives via their husbands (Gangakhedkar et al., 1997, & Newmann et al., 2000, as cited in Solomon et al., 2010). Since heterosexual transmission was the primary driver of HIV in India, most prevention programs have targeted these individuals; recent evidence suggests that the HIV epidemic is declining among these populations (Arora et al., 2008, & Kumar et al., 2006, as cited in Solomon et al., 2010). By contrast, MSM remain the only high-risk population in India where HIV prevalence is increasing (National AIDS Control Organization, 2007, as cited in Solomon et al., 2010). There are an estimated 2.35 million sexually active MSM in India (National AIDS Control Organization, 2006, as cited in Solomon et al., 2010) ; historically, MSM in India have been a hidden population because of Section 377 of the Indian Penal Code which criminalizes anal intercourse and Indian culture's requirement of opposite sex marriage (Herget, 2006, as cited in Solomon et al., 2010). Consequently, few studies of HIV epidemiology among MSM in India exist. Most derive from urban sexually transmitted infection (STIs) clinics and few have characterized MSM in rural/semi-urban areas. In a landmark decision on July 2, 2009, the Delhi High Court repealed Section 377 (Nilajana, 2009, as cited in Solomon et al., 2010). While this allows more freedoms for MSM, it is likely that stigma and discrimination will continue and may even increase as MSM become more visible in Indian society. A study conducted across 18 different cities (classified in six urban sites, six semi-urban, six rural) in the south Indian state of Tamil Nadu showed HIV and syphilis prevalence of 9 and 8% respectively among not married MSM and of 14 and 11% among married MSM; HIV infections were not restricted to the large metropolitan cities because several semi-urban and rural settings also had HIV-infected MSM (Solomon et al., 2010).

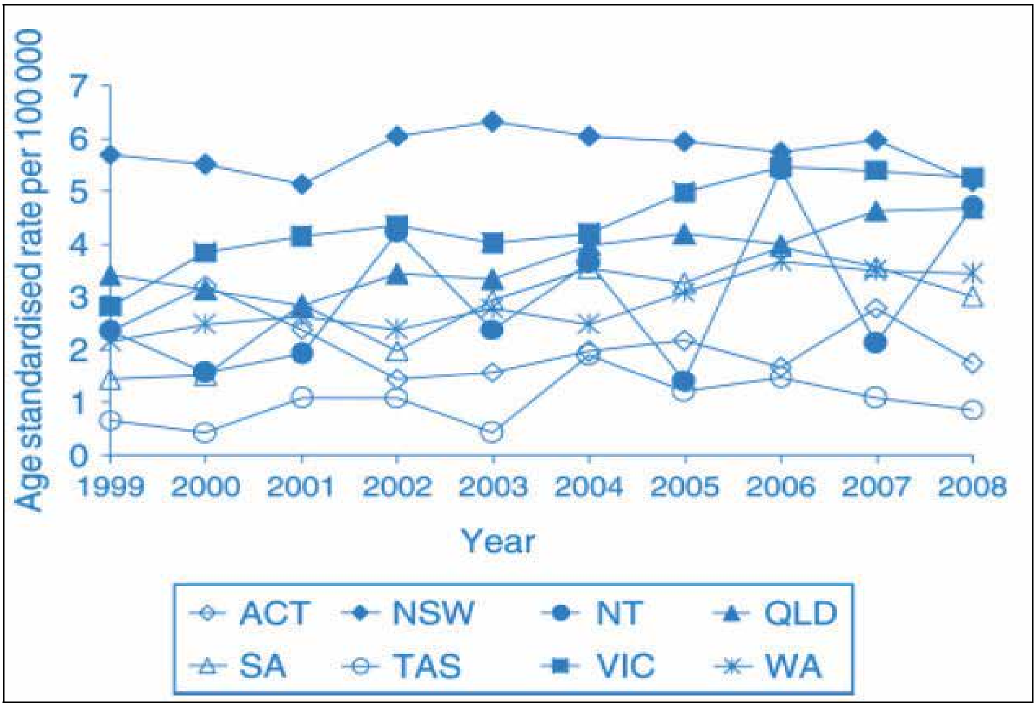


Fig. 1. Newly diagnosed HIV infection, 1999–2008, by year and state and territory, Australia. (Source: de Wit et al., 2010)

2.2.2 Sub-Saharan Africa

In sub-Saharan Africa the scenario is similar to that of India because research on MSM was neglected for a long time; most African countries do not include MSM in their HIV surveillance systems. Criminalisation of male-to-male sex, social denial and difficulty in reaching this population were the explanations given for the weak involvement of researchers in this area (Dahoma et al., 2011). However, recently, the assumption that this population could be significantly vulnerable to HIV and STIs has encouraged a growing number of epidemiological studies. Although heterosexual contact remains the main mode of HIV transmission in sub-Saharan Africa, HIV prevalence there has been described to be higher within MSM population than in the general population. In the first epidemiological study conducted on MSM in Senegal, HIV prevalence was 21.5% compared with less than 1% for the general Senegalese population. The high frequency of sexual contact with women reported by this study's participants also suggests that a sexual bridge between MSM and women could contribute to the expansion of the HIV epidemic (Wade et al., 2005). High levels of HIV infection also were reported among MSM in Zanzibar where a survey conducted from April through June 2007 among 509 MSM participants showed HIV prevalence of 12.3% (65/509) (Dahoma et al., 2011). Moreover, the study showed high levels of bisexual activity (70.8% of participants) which in combination with high risk sexual behaviour (i.e. inconsistent condom use) raise concerns about the potential bridging role of MSM in transmitting HIV infection into the general population (Dahoma et al., 2011). In Nigeria results from a survey conducted in Lagos, Kano and Cross River states during 2007 showed that MSM are highly affected by HIV and that condom use in anal sex is low in commercial and non-commercial partnerships. HIV is more established among MSM in Lagos, there were wide variations in HIV prevalence with estimates ranging from 17.4%, to 9.3% and 1.1% in Lagos, Kano and Cross River respectively (Merrigan et al., 2011). Moreover, as previously observed in other African countries, significant proportions of Nigerian MSM had unprotected sex with female partners indicating a potential bridge for HIV transmission between MSM and the general population (Merrigan et al., 2011).

These data suggest that prevention interventions should be systematically implemented within the high-risk group of MSM in the African context.

2.2.3 China

In China, men who have sex with men have now become one of the priority populations for prevention and control of HIV pandemic because MSM transmission of HIV is a critical and growing public health problem. The estimated number of people living with HIV in China has also increased markedly from 400.000 in 2007 to 740.000 in 2009. Among new HIV infections reported in China, the proportion of MSM transmission has risen dramatically. According to pooled prevalence estimates from meta-analyses (Chow et al., 2011), the overall national HIV prevalence among MSM in China has increased from 1.3% during 2003-2004 to 2.4% during 2005-2006 and then to 4.7% during 2007-2008; moreover, a significantly positive correlation was observed between HIV and syphilis prevalence among Chinese MSM during 2003–2008 countrywide.

With substantial increases in HIV and syphilis infection, Chow et al. (2011) also observed an increase in prevalent HIV-syphilis co-infections (1.4% during 2005-2006 to 2.7% during 2007-2008) (Figure 2). In 2009, 32.5% of 48.000 new HIV infections was attributed to male homosexual exposure, which is almost triple the reported level in 2007 (12.2% of 50.000 new cases) (UNGASS, 2010, as cited in Chow et al., 2011). In Shenyang city, the capital of Liaoning Province in northeast China, MSM transmission accounted for 39.6% of new infections in 2009. Epidemiological studies from 2006 suggested that HIV infection among MSM in Shenyang was increasing rapidly, with an HIV incidence density of 5.4/100 person-years (Xu et al., 2010). A 3-year prospective open cohort follow-up survey (from 2007 to 2009) that recruited a total of 1282 MSM (229 in 2007, 598 in 2008 and 455 in 2009) revealed a significant increase in incidence density of HIV from 4.7/100 person-years to 10.2 /100 person years (Table 1). Furthermore, due to a higher transmission probability of HIV associated with penile-anal intercourse and levels of risk related behaviour, Chinese MSM have an approximately 45-fold higher risk of acquiring HIV than other males in the general population (Zhang et al., 2011).

Type of recruited MSM cohorts	*Infection rate*	*Incidence density (95% CI)*
Recruited cohort in 2007	5.7 %	4.7/100 person-years
Recruited cohort in 2008	6.2 %	6.1/100 person-years
Recruited cohort in 2009	8.1 %	10.2/100 person-years

Table 1. China. Shenyang city : new infection rates and incidence density of HIV. (Source: Zhang et al., 2011)

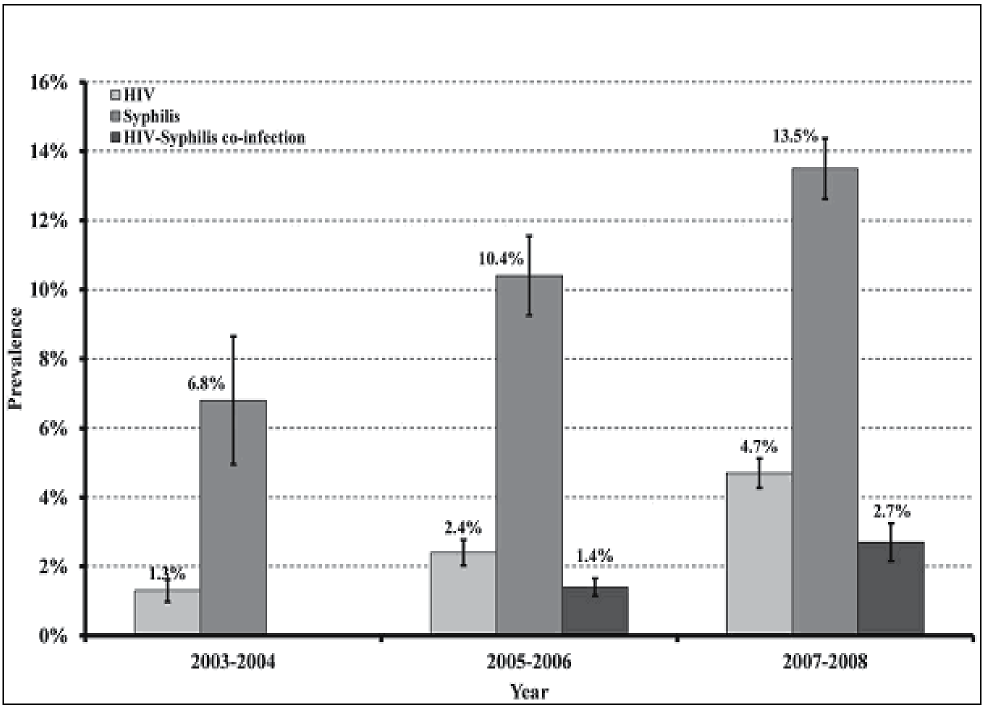


Fig. 2. China : estimated prevalence of HIV, syphilis infection and HIV-syphilis co-infection among men who have sex with men during 2003–2008. (Source : Chow et al., 2011)

3. Treponema pallidum / Syphilis

Syphilis. The name of this widespread and ancient disease is familiar to health care providers worldwide, it was coined by the Italian physician Girolamo Fracastoro in his tale of the shepherd Syphilus, who contracted the disease after insulting the pagan Sun god.

Syphilis is a genital ulcerative disease caused by the spirochaete *Treponema pallidum* subsp. *pallidum* (Fenton et al., 2008), identified in 1905 by Schaudinn & Hoffmann (Waugh, 2005, as cited in Fenton et al., 2008). The disease is treatable with standard antibiotics (Pao et al., 2002, as cited in Fenton et al., 2008), but in the absence of therapy, gives rise to four sequencial clinical stages commonly known as primary, secondary, latent (divided into early and late stages) and tertiary syphilis (Table 2) (Botham et al., 2006). In the initial stages, the infection is highly transmissible and often presents as ulcerative genital lesions. Secondary syphilis results from the multiplication and dissemination of Treponemes throughout the body. The most common manifestation is a disseminated mucocutaneous rash.

The secondary stage occurs up to 6 months after the healing of the primary lesion and lasts for several weeks or months and may reoccur in approximately 25% of untreated patients (Stamm, 1998, as cited in Fenton et al., 2008). The oral cavity is the most common extra-genital site of infection; oral chancres are observed in about 4–12% of patients with primary syphilis and occur at the site of penetration of the organism into the mucosa. Anatomical sites commonly affected are especially the tongue, gingiva, soft palate and lips (Ficarra & Carlos, 2009) (Figure 3). The tertiary or late stage of syphilis is rarely seen today in the era of effective and prevalent antibiotic therapy. Treponemes invade the central nervous or cardiovascular systems, eyes, skin, and other internal organs, producing damage as a result of their invasive properties and inflammation (Fenton et al., 2008). With minor differences, syphilis generally presents similarly in HIV-positive and HIV-negative patients. In primary syphilis, HIV-infected patients may present with more than one chancre (up to 70% of patients) and with larger and deeper lesions. About a quarter of HIV-infected patients present with concomitant lesions of both primary and secondary stages of syphilis at the time of diagnosis. Approximately one-third of patients with early syphilis have invasion of Treponemes in the cerebrospinal fluid (CSF), regardless of their HIV status. However, by contrast with HIV-uninfected patients, most of the new cases of clinical neurosyphilis in HIV-infected individuals are identified early at the initial presentation, suggesting that HIV infection may be associated with an increased risk of developing neurological complications (Zetola & Klausner, 2007).

3.1 The current scenario in developed countries

Rates of syphilis in many developed cities worldwide reached their lowest recorded levels towards the end of the 1990s. Since 1997, however, outbreaks of syphilis among men who have sex with men have been reported in North America and Europe, and those with HIV infection have been disproportionately affected (CDC, 2002). Australia has two distinct patterns of infectious syphilis: a substantially declining occurrence in Indigenous remote communities and an increasing incidence in males residing in urban and regional areas (Ward et al., 2011).

3.1.1 European Union and European Free Trade Association countries

In the early 1990s rates of syphilis infection in Western Europe (WE) were at historically low levels, these decreases were accompanied by marked reductions in the incidence of congenital syphilis and tertiary disease. But **syphilis has re-emerged in western Europe**

since 2000 and a large increase and outbreaks affecting major urban centres in Europe, mainly in MSM, have been reported (Figure 4) (Sullivan et al., 2009). **Changes in sexual behaviour have facilitated the spread of syphilis especially among men who have sex with men and improved surveillance systems and case detection have lead to an increase in the reported numbers of cases.** The first of several outbreaks of syphilis in MSM was observed in Hamburg (Germany) in 1996/1997, and syphilis incidence began to increase nationwide in 2001 (Marcus & Hamouda, 2005); in Germany, for syphilis cases in MSM, 7% in the year 2002 were classified as re-infections.

INFECTIOUS	NON-INFECTIOUS
Primary One or more ano-genital or oral ulcers (chancres) are present, which may vary considerably in appearance. Serological tests are reactive. **Secondary** Skin spots or rashes are present, particularly on the trunk, palms and soles,often with generalised lymphadenopathy. The primary chancre may still be present. Non-treponemal (RPR, VDRL) titre 1:4. **Latent** No symptoms of syphilis are present, but treponemal (TPPA, FTA-Abs) tests are reactive and the non-treponemal (RPR, VDRL) tests have increased two titres. **Early latent** Disease acquired within the past two years.	**Late latent** Disease acquired more than two years ago or at an unknown time. **Neurosyphilis** Syphilis of any stage with clinical symptoms of neurosyphilis or raised CSF protein or WCC in the absence of other known causes of these abnormalities. **Tertiary** Characteristic abnormalities of the cardiovascular, skin, bone or other systems. **Congenital syphilis** A condition affecting an infant whose mother had untreated or inadequately treated syphilis at delivery.

Table 2. Classification of syphilis. (Source: Botham et al., 2006)

This proportion increased to 16% in 2004, in seven cities (Frankfurt, Cologne, Berlin, Mannheim, Munich, Leipniz, Hamburg) with > 100.000 population, with the highest incidence of syphilis in 2004 (>10/100.000), 25% of reported cases in MSM were classified as re-infections. This datum indicates an intense circulation of syphilis in restricted subgroups of MSM in the largest cities (Marcus & Hamouda, 2005). In Denmark an outbreak of syphilis began in 2003 and continued in 2004. Data from the national surveillance system showed that most cases are in men who have sex with men, and that a large proportion of these patients are also HIV positive. The majority of cases were reported from Copenhagen municipality; the outbreak did not seem to be affecting the age group under 20 years. Probably, most of the MSM found with both syphilis and HIV were already HIV positive when they acquired syphilis infection (Cowan, 2004). In the United Kingdom, between 1997 and 2002, diagnoses of primary, secondary, and early latent syphilis made at genitourinary medicine clinics increased by 213% in heterosexual males, 1412% in men who have sex with men (MSM), and 22% in females. These increases were initially observed in larger cities and then progressed to suburban and rural settings (Simms et al., 2005). In 2004, among MSM/bisexual men an alarming increase of 212% as regards 2003 has been observed in

Scotland; most of the diagnoses (91%) were made in the Glasgow and Edinburgh clinics (Figure 5) (Wallace et al., 2005). The Swedish Surveillance system of syphilis is population-based. All health care specialists have the responsibility to notify diagnosed syphilis directly to the national surveillance system SmiNet. Also all laboratories diagnosing syphilis have the responsibility to notify to the SmiNet. From the data obtained through the surveillance system seems that the major contributor to the recent rise in syphilis cases in Sweden is infections among MSM (Figure 6) (Velicko et al., 2008). Among MSM who acquired a syphilis infection in Sweden in recent years, the majority acquired it in big cities, especially Stockholm (up to 96% of all reported cases among MSM); syphilis incidence among MSM was 20-28 times that of males in general in Sweden.

Fig. 3. Secondary oral syphilis: mucous patches covered by *grayish, white* pseudomembranes of the lower vestibular mucosa. (Source : Ficarra & Carlos, 2009).

It can be assumed that MSM to a larger extent choose to live in big cities since they assure more anonymity and less stigmatisation for MSM (Velicko et al., 2008). In contrast to WE countries, in Central and Eastern Europe between 1998 and 2007 there was a general decline in the number of reported syphilis cases (Savage et al., 2009). During the same period the sex ratio was relatively stable at around 1:1, with the exception of Slovenia where the male: female ratio rose to almost 5:1 in 2007 and 39% of male cases were reported to be MSM (Savage et al., 2009). Recent increases in overall early syphilis reported rates in Slovenia (in 2008 increase of 130% compared with 2007), with great majority of cases occurring in men (94%) and among cases in men a substantial proportion of cases among men known to have sex with men (44%) and among those almost one in three with a foreign MSM partner within three months preceding syphilis diagnosis, suggest recent increase in unsafe sexual behaviour among MSM and sexual mixing of Slovenian MSM with MSM abroad (Klavs et al., 2009). Most worrying, two in three early syphilis cases reported in 2008 among MSM were men with known HIV infection, indicating unsafe sexual behaviour among HIV-positive MSM aware of their infection. In the Czech Republic, most cases of syphilis, total of 18.000 were diagnosed after the World War II during the mass screening of the population aged 15 to 40 years in 1951. In the following years, the incidence of syphilis decreased considerably; the absolute number of reported syphilis cases has increased rapidly during the 1990s as a result of geopolitical and socioeconomic changes (Kuklova et al., 2011). The syphilis rate has risen eightfold from 1.6 per 100.000 inhabitants in 1990 to 13.5 per 100.000 inhabitants in 2001 due to high proportion of syphilis cases among immigrants coming mainly from the Ukraine, Moldavia, the Russian Federation and Georgia. After reaching a peak in 2001, the incidence of syphilis has been decreasing but, after 2006, the incidence has increased again. Men, in particular MSM, prevail over women among the infected persons (Kuklova et al., 2011).

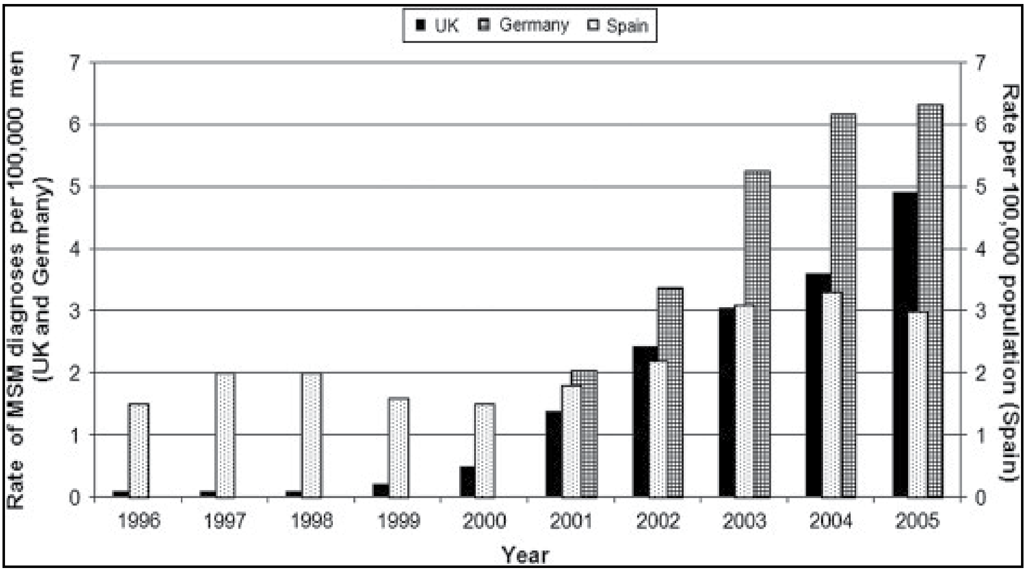


Fig. 4. Primary and secondary syphilis rates in United Kingdom and Germany (rate of diagnoses among MSM per 100,000 men) and Spain (rate per 100,000 population), 1996-2005. (Source: Sullivan et al., 2009).

3.1.2 United States / Canada

Although syphilis is uncommon in much of the US population, several groups are disproportionately affected, including African Americans, Hispanics, and men who have sex with men (MSM).

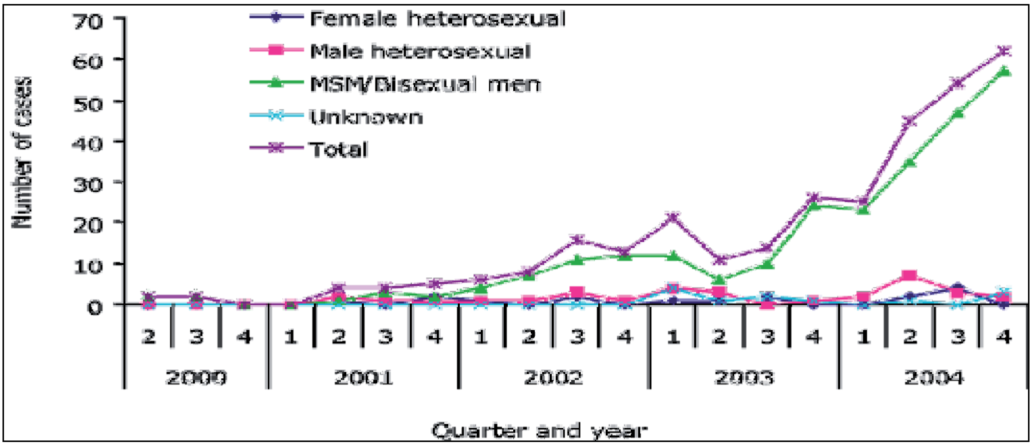


Fig. 5. Number of cases of infectious syphilis diagnosed at STI clinics in Scotland by sexual orientation and quarterly totals, 2000-2004. The majority of cases were diagnosed at sexually transmitted infections (STIs) clinics in Edinburgh (45%) and Glasgow (42%). (Source : Wallace et al., 2005)

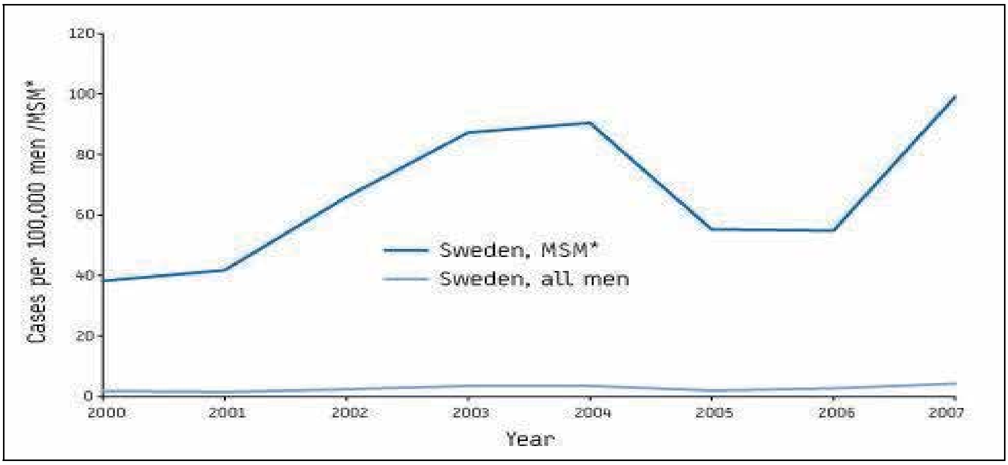


Fig. 6. Estimated syphilis incidence among MSM and observed syphilis incidence among all males, notified cases in Sweden, 2000-2007. (Source: Velicko et al., 2008)

After steady decreases for >1 decade, rates of syphilis in the United States reached their lowest point during 2000, when the rate of primary and secondary syphilis was 2.1 cases per 100,000 persons. By 2005, the rate of primary and secondary syphilis increased to 3.0 cases per 100,000 persons, representing an increase from 5976 to 8724 cases, of which 86% occurred in men. This increase in the rate ratio of male to female patients (from 1.2 in 1996 to 5.7 in 2005) was a reflection of the disproportionate burden of disease among MSM, who accounted for ~65% of all persons with primary and secondary syphilis (Zetola & Klausner, 2007). Large metropolitan communities of MSM have been particularly affected by concurrent epidemics of syphilis and HIV infection; genital ulcers bleed frequently during sexual intercourse, resulting in potential increases in HIV infectiousness. In California more of 60% of MSM with syphilis are HIV infected, and it is estimated that, in major cities, 20-50% of MSM with syphilis have also HIV infection. The increase in the rate of syphilis and other sexually transmitted diseases (STDs) among MSM suggests a decrease in safer sex practices. In this regard, the success of HAART, the use of the Internet to meet sex partners, the increased frequency of serosorting (i.e., finding sex partners with the same HIV serostatus), and the increase in recreational drug use, both illicit (e.g., crystal methamphetamine) and prescribed (e.g., sildenafil citrate), have all likely contributed to increases in the rate of syphilis. In addition, the idea that oral sex is "safer" sex and rarely associated with HIV transmission may explain the role of oral sex in syphilis transmission (Zetola & Klausner, 2007). Recently, considering the high incidence of asymptomatic syphilis in HIV positive MSM, has been stressed the importance of frequent routine syphilis testing in this group of population. In USA, syphilis cases occurring among MSM increased from 441 in 2000 to 4424 in 2003; the estimated percentage of all reported cases occurring among MSM in 2003 was 62% (Table 3) (Heffelfinger et al., 2007). During 2000 through 2003, there was an estimated increase of 2298 cases among white MSM; in contrast, cases among heterosexual men and women decreased 50% (from 5532 to 2753) between 2000 and 2003.

In Canada, a review of infectious syphilis cases carried out in Ottawa, from 2001-2006, showed that rates of syphilis have risen more than tenfold (Leber et al., 2008) and that the majority of cases (83.5%) occurring among men who have sex with men. These individuals differed from the general MSM population residing in Ottawa in their being older, more likely to be HIV positive, and more sexually promiscuous; inconsistent condom use by MSM engaged in either oral or anal sex was pervasive. Visceral manifestations of syphilis, including neurosyphilis, were more common in persons co-infected with HIV. Multiple sexual partners, unprotected oral sex, and increased age among MSM were the predominant risk factors contributing to this syphilis epidemic (Leber et al., 2008). Co-infection with HIV modified the clinical presentation of syphilis, necessitating a more intensive diagnostic and therapeutic approach. The interconnection of urban sexual networks has likely contributed to the dynamics of local syphilis transmission and suggests that effective interventions will require a coordinated national approach (Leber et al., 2008).

Year	Total No. of cases	Cases among men	Cases among women	Male-Female Case ratio	Cases among MSM	Cases among heterosexual men and women [a]
2000	5973	3528	2445	1.44	441	5532
2001	6100	4132	1968	2.10	1653	4447
2002	6862	5268	1594	3.30	3257	3605
2003	7177	5959	1218	4.89	4424	2753

a. Calculated by subtracting estimated cases among MSM from the total number of cases.

Table 3. Numbers of Primary and Secondary Syphilis Cases Among Different Groups, Along With Male–Female Case Ratios: United States, 2000–2003. (Source: Heffelfinger et al., 2007).

3.1.3 Australia

Syphilis appeared to be well controlled in large urban centres in Australia during the 1990s, probably as a result of HIV prevention efforts; an overall decrease in syphilis rate was seen until 2001. Since 2001, enhanced syphilis surveillance has been undertaken in south-eastern Sydney for all new laboratory syphilis notifications. The New South Wales (NSW) surveillance classifications for newly diagnosed adult cases are showed in Table 2. However, in 2002 the overall rates of syphilis increased by 14.1%, possibly related to public health activities such as active testing for STIs at selected sex-on-premise venues (Victoria), a targeted syphilis campaign coinciding with the Gay Games in Sidney in 2002 (New South Wales), and the establishment of a syphilis registry (Queensland) (Yohannes et al., 2004, as cited in Botham et al., 2006). In particular, inner Sidney has recently experienced a rapid increase in infectious syphilis, similar to other developed cities with large gay male populations. During 2001-04, 1275 syphilis notifications were received ; of these notifications 1112 were classified as 361 infectious syphilis, 221 non-infectious syphilis and 530 treated syphilis. The majority of infectious syphilis diagnoses were made in the earlier primary or secondary stages. The enhanced syphilis surveillance program showed infectious syphilis is affecting a defined population in Sidney: men (31:1 male-to-female ratio for 2004), English speaking and Australian born (Botham et al., 2006). The male-to-female ratio for 2004 of 31:1 strongly suggests transmission between men (Botham et al., 2006). The public health response to this epidemic has included work by the Sexually Transmitted Infections in Gay Men Action Group (STIGMA). This group was formed in 2000 and initiated the project for enhanced syphilis surveillance and also developed and disseminated guidelines entitled *Sexually Transmitted Infection Testing Guidelines for Men who have Sex with Men*. These guidelines recommend annual syphilis serology for all men who have had sex with another man in the previous year; this recommendation, promoted to clinicians and the gay community, may have contributed to the overall increase in syphilis notifications. Data of literature (Jin et al., 2009) showed that in Sidney STIs peak among MSM occurred in the 35-44 year age group, suggesting a decline in sexual "safety" among older gay men, and that syphilis incidence was almost 10-fold higher in HIV-positive than in HIV-negative MSM and it was not related to a CD4 count below 350 µL ; UAI with HIV positive partners was of particular importance in the transmission of syphilis. Within Australia, syphilis has also increased largely in gay men in the Queensland and Victoria (Bourke & Schmidt, 2009).

3.2 The current scenario in developing/transitional countries

Infectious diseases continue to cause an enormous burden of death and disability in developing countries; a relevant role is represented from sexually transmitted infections. Even if the last years have seen a re-emergence of international interest in the role of MSM in STDs epidemics globally, data on their incidence and prevalence, in particular of syphilis, are still poor in most of the developing countries. This is most evident in sub-Saharan Africa, north Africa and the Caribbean, where stigma around homosexuality is still considered appalling. However, even in Asia and Latin America, studies conducted usually lack a clear understanding of these populations and their diverse behaviours (Càceres et al., 2008).

3.2.1 India

Reports from sexually transmitted disease clinics and single health care institutions in India suggest an increase in genital ulcerative sexually transmitted infections such as syphilis. The little information available on rates of syphilis in India show an incidence rate of 5.4 per 100 person years in a sexually transmitted disease clinic in India from 1993-2000 and prevalence rates as high as 21.9% amongst a convenience sample of long-distance truck drivers in 2000 (Gawande et al., 2000, & Reynolds et al., 2006, as cited in Schneider et al., 2010). A recent study (Schneider et al., 2010) carried out in Andhra Pradesh in the general population (6.235 men and 6.382 women study participants) showed prevalence rates of 2.08% for men and 1.42% for woman; nearly one in four persons surveyed that were seropositive for syphilis were also HIV infected. 85 men, that declared to have sex with men > 6 months ago, showed a syphilis prevalence of 2.4% while other 47, that declared own homosexuality within last 6 months, showed a prevalence of 6.4%. The strong association between syphilis and HIV seroprevalence in this population-based study suggests that acceleration of direct linkages between syphilis testing and HIV counselling and testing would be useful in enhancing the control of STIs and HIV in India; particular attention should be paid to homosexuals men of India that have one of the highest rates of HIV prevalence amongst groups at high risk of sexual transmission of HIV.

3.2.2 Sub Saharan Africa

In 1999, WHO estimated that were 4 million cases of syphilis among adults in sub-Saharan Africa. In Africa syphilis is endemic and the incidence is high also among pregnant woman; this is a worrying fact because maternal syphilis has a severe impact on pregnancy outcome and there are significant challengens to implementing syphilis screening programmes in sub-Saharan Africa (Watson-Jones et al., 2005). Many prevalence studies have been conducted in the general population but unfortunately, data on the prevalence of syphilis in MSM are scanty, likely because of stigma and discrimination against homosexuality. Some studies conducted among blood donors have shown a variable prevalence among different sub - Saharan geographical areas : in Burkina-Faso a study (Nagalo et al., 2011) showed that seroprevalence of syphilis (3.96%) was lower than the prevalence of among blood donors in Tanzania (4.6% among HIV negative, 12.1% among HIV positive) (Matee et al., 2006) but higher than the 1.1% observed in Nigeria (Buseri et al., 2009), where recently no syphilis was detected among 879 examined MSM (Merrigan et al., 2011) suggesting a lower prevalence of syphilis in Nigerian population compared to other African populations. A recent study carried out in Ghana in a sample of 1366 prison inmates, including 403 homosexuals men, revealed seroprevalence syphilis rates of 13.7 % and 23.3% in heterosexual and homosexual respectively and syphilis seropositivity was significantly associated with HIV seropositivity (Adjei et al., 2008). These results confirm data previously reported and emphasize the need for preventive interventions in areas such as prisons that are known to be high-risk environments for the spread of bloodborne and sexually transmitted infections.

3.2.3 China

Syphilis was nearly eradicated in China in 1964 (Cohen et al., 1996, as cited in Chow et al., 2011) but recently it has returned to become one of the top five most reported notifiable diseases. Has been documented a 30-fold increase in syphilis diagnoses from 0.2 cases per 100.000 individuals in 1989 to 6.5 cases per 100.000 individuals in 1999 and a further 3-fold increase over the following decade to 22 cases per 100.000 individuals in 2008 (Tucker et al., 2010, as cited in Chow et al., 2011). Cross-sectional studies have found a high prevalence of syphilis among men who have sex with men in China. Sample representativeness remains one of the challenges in effective STDs surveillance and prevention targeting men who have sex with men worldwide. Although convenience samples are widely used in studies of MSM, previous studies suggested that these samples might not be representative of the broader MSM population. This issue becomes even more critical in many developing countries where needed resources for conducting probability sampling are limited. A study carried out in Beijing during 2009 among four MSM samples using different recruitment methods showed rates of Syphilis infection of 21.8%, 36.2%, 11.8%, and 13.8%; rates of inconsistent condom use were 57%, 52%, 58%, and 38%. Significant differences were found in various socio-demographic characteristics (e.g., age, migration history, education, income, and places of employment) and risk behaviours (e.g., age at first sex, number of sex partners, involvement in commercial sex, and substance use) among samples recruited by different sampling methods (peer outreach, informal social network, Internet, and venue-based) (Guo et al., 2011). Other recent studies (Xu et al., 2011; Zhang et al., 2011) carried out in Shenyang city (in 2009) and Liaoning (from April 2008 to January 2009) showed a syphilis prevalence of 14.9% and 5% respectively. These variable results of seroprevalence confirm the importance of using multiple sampling methods to reach MSM from diverse backgrounds and in different social segments and to improve the representativeness of the MSM samples when the use of probability sampling approach is not feasible.

4. Conclusion

Men who have sex with men, among all races and ethnicities, remain at great risk for HIV infection and syphilis; behavioural prevention remains central to reduce their transmission and can be the key to this process. Sex is life, but we have to be conscious of the fact that safe sex is the first premise. Sexually transmitted diseases (STDs) are among the first ten causes of unpleasant diseases in young adult males in developing countries and the second major cause of unpleasant diseases in young adult women worldwide. In general, STDs are epidemics and present enormous health and economic consequences. The surest way to avoid transmission of sexually transmitted diseases, including syphilis, is to abstain from sexual contact or to be in a long-term mutually monogamous relationship with a partner who has been tested and is known to be uninfected. Avoiding alcohol and drug use may also help prevent transmission of syphilis because these activities may lead to risky sexual behavior (Zetola & Klausner), (Mayer et al. 2010). It is important that sex partners talk to each other about their HIV status and history of other STDs so that preventive action can be taken. An adequate screening for re-emerging STDs should be done on a routine basis in every part of the world, in particular among highly vulnerable subjects at risk for HIV and syphilis as homosexual men; individuals unaware of their infection are a priority for being tested and diagnosed to enable them to take advantage of care services and receive counselling to prevent further spread of HIV and *T. pallidum*. Partner selection based on perceived serostatus could be used as a strategy for risk reduction among MSM but it strategy has some limits since those most recently infected may be most infectious and least likely to know of their status, reliance on partner’s awareness and disclosure of their own serostatus may be a risky strategy. Individuals who are infected with STDs are 5-10 times more likely than uninfected individuals to acquire or to transmit HIV through sexual contact. The breaking of the genital tract or trauma of the rectal mucosa creates a portal of entry for HIV, and HIV positive individuals co-infected with other STDs, especially ulcerative infections such as syphilis, are more likely to shed HIV in their genital secretions. The most effective method available for protection against STDs and HIV is the correct and consistent use of condom that is increasing everywhere except in developing countries. In particular, high risk sexual behaviours reported in Africa and in India among MSM and high levels of bisexual activity are of grave concern. Most worrisome is that condom use messages for preventing HIV and STDs do not appear to have penetrated this population given the low proportion of MSM using condom ; three most often repeated reasons for not using condoms with regular partners include non availability, feeling the partner is safe and feeling that condoms do not give sexual pleasure (Kumta et al., 2010). Interestingly, a study carried out in the south-eastern part of India (Andhra Pradesh) showed that one of the primary reasons for low condom usage was that many married MSM did not disclose their sexual orientation to their wives and therefore struggled to explain the necessity of condom use (Gutierrez et al., 2010). This datum suggests that in some parts of the developing world prevention work is handicapped unless it addresses broader issues surrounding sexual identity, stigma and discrimination. Internet is commonly used by MSM to find sexual partners. The net is a relatively safe space where sexuality can be addressed in a frank and honest way because the web allows anonymity and confidentially. In a context where homosexuality is penalised, blackmail and police harassment are common and the media are homophobic, internet may be a way to support and inform men who have sex with men about HIV/STDs and other issues which men ask about (Pebody, 2010). For example, in

Romania, Kovacs used a sexual networking site to build a support network of HIV positive MSM (an extremely marginalised and, until then, invisible group in Romania). In another example, in Thailand, Mplus Organisation works with various groups including men selling sex, transgender people and MSM; the organisation lets their existing contacts know that staff will be present on various social networking sites and MSM Messenger at specific times, where they are available to interact and answer questions (Pebody, 2010). These data suggest that Internet is a double edged sword because on the one hand promotes sexual encounters and promiscuity, on the other hand can help prevent the spread of HIV and other sexually transmitted infections such as syphilis. Interestingly, during a recent syphilis outbreak that involved 7 MSM in San Francisco who met in an online chat, electronic notification of the cluster was provided to hundreds of chat room users. Despite very limited contact information, the Department of Public Health was able to notify and treat 42% of the named sex partners (Klausner et al., 2000). The major challenge for the future is to reduce the rates of new infections in a context of increasingly diverse and complex prevention responses. A key requisite for this development is the occurrence of transformations that tackle homophobic violence and ensure the respect for the human rights of gay men, transgenders and other MSM.

In conclusion, more research is need to further clarify which strategies (i.e. reducing unprotected anal sex, condom distribution, having oral sex instead of anal sex, reducing number of partners, avoiding serodiscordant partners, strategic positioning or reducing anal sex even with condom use) are most effective in reducing sexually transmitted infections transmission among MSM, the messages most effective in promoting preventive behaviours, and the methods and settings in which these messages can be most effectively delivered.

5. References

Adjei, A.A., Armah, H.B., Gbagbo, F., Ampofo, W.K., Boamah, I., Adu-Gyamfi, C., Asare, I., Hesse, I.F. & Mensah, G. (2008). *Correlates of HIV, HBV, HCV and syphilis infections among prison inmates and officers in Ghana: A national multicenter study.* BMC Infect Dis, 8:33.

Botham, S.J., Ressler, K.A., Bourne, C. & Ferson, M.J. (2006). *Epidemic infectious syphilis in inner Sydney--strengthening enhanced surveillance.* Aust N Z J Public Health, 30(6):529-33.

Boulos, D., Yan, P., Schanzer, D., Remis, R.S. & Archibald, C.P. (2006). *Estimates of HIV prevalence and incidence in Canada, 2005.* Can Commun Dis Rep 32(15):165-74

Bourke, S. & Schmidt, T. (2009). *Sexually transmissible infections - Old enemies and a new friend.* Aust Fam Phisician, 38(6):373.

Buseri, F.I., Muhibi, M.A. & Jeremiah, Z.A. (2009). *Sero-epidemiology of transfusion-transmissible infectious diseases among blood donors in Osogbo, south-west Nigeria.* Blood Transfus, 7(4):293-9.

Cáceres, C.F., Konda, K., Segura, E.R. & Lyerla, R. (2008). *Epidemiology of male same-sex behaviour and associated sexual health indicators in low- and middle-income countries: 2003-2007 estimates.* Sex Transm Infect, 84 Suppl 1:i49-i56.

CDC. (2002). *Primary and secondary syphilis among men who have sex with men- New York city, 2001.* MMWR, 51 : 853-6

CDC. (2004). *Trends in primary and secondary syphilis and HIV infections in men who have sex with men – San Francisco and Los Angeles, California, 1998-2002.* MMWR, 53 : 575-78

CDC).(2005). *HIV/AIDS surveillance report, 2004.* Available from http://www.cdc.gov/hiv/stats/hasrlink.htm

CDC. (2006). *Twenty-five years of HIV/AIDS- United States, 1981-2006.* MMWR, 55 (21): 585-89

CDC. (2007). *HIV/AIDS Surveillance Report 2006.* Atlanta (GA) : U.S. Department of Health and Human Services, CDC and Prevention 181- 46

CDC. (2008a). *Trends in HIV/AIDS diagnoses among men who have sex with men.* MMWR, 57:681-86

CDC. (2008b). *Subpopulation estimates from the HIV incidence surveillance System- United States 2006.* MMWR, 57: 985-89

CDC. (2011). *Increase in newly diagnosed HIV infections among young black men who have sex with men- Milwaukee County, Wisconsin, 1999-2008.* MMWR, 60:99-102

Chow, E.P., Wilson, D.P. & Zhang, L. (2011*). HIV and Syphilis Co-Infection Increasing among Men Who Have Sex with Men in China: A Systematic Review and Meta-Analysis.* 6(8):e22768. PLos One, Epub 2011 Aug 15.

Couturier, E., Michel, A., Janier, M., Dupin, N. & Semaille, C. *Syphilis surveillance in France, 2000-200*3. Euro Surveill. 2004 ; 9 (12) : pii = 493

Cowan, S. (2004).*Syphilis in Denmark-Outbreak among MSM in Copenhagen, 2003-2004.* Euro Surveill, 9 (12) :25-7.

Dahoma, M., Johnston, L.G., Holman, A., Miller, L.A., Mussa, M., Othman, A., Khatib, A., Issa, R., Kendall, C. & Kim, A.A. (2011). *HIV and related risk behaviour among men who have sex with men in Zanzibar, Tanzania: results of a behavioural surveillance survey.* AIDS Behav, 15(1):186-92.

Dandona, L., Dandona, R., Gutierrez, J.P., Kumar, G.A., McPherson, S. & Bertozzi, S.M. (2005). *Sex behaviour of men who have sex with men and risk of HIV in Andhra Pradesh, India.* AIDS, 19(6):611-9.

de Wit, J.B., Prestage, G.P. & Duffin, I.R. (2010). *Gay men: current challenges and emerging approaches in HIV prevention.* N S W Public Health Bull, 21(3-4):65-68.

Downs, A.M., Heisterkamp, S.H., Ravà, L., Houweling, H., Jager, J.C. & Hamers,.F.F. (2000). *Back-calculation by birth cohort, incorporating age- specific disease progression, pre-AIDS mortality and change in European AIDS case definition. European Union Concerted Action on Multinational AIDS Scenarios.* AIDS; 14(14):2179-89.

Elford, J., Jeannin, A., Spencer, B., Gervasoni, J.P., van de Laar, M.J. & Dubois-Arber, F. (2009). *HIV and STI behavioural surveillance among men who have sex with men in Europe.* Euro Surveill, 14(47). pii: 19414

Fenton, K.A., Breban, R., Vardavas, R., Okano, J.T., Martin, T., Aral, S.& Blower, S. (2008). *Infectious syphilis in high-income settings in the 21st century.* Lancet Infect Dis 8 (4) :244-53

Ficarra, G. & Carlos, R. (2009). *Syphilis: the renaissance of an old disease with oral implications.* Head Neck Pathol, 3(3):195-206.

Guo, Y., Li, X., Fang, X., Lin, X., Song, Y., Jiang, S. & Stanton, B. (2011). *A comparison of four sampling methods among men having sex with men in China: implications for HIV/STD surveillance and prevention.* AIDS Care, Epub ahead of print

Gutierrez, J.P., McPherson, S., Fakoya, A., Matheou, A. & Bertozzi, S.M. (2010). *Community-based prevention leads to an increase in condom use and a reduction in sexually transmitted infections (STIs) among men who have sex with men (MSM) and female sex workers (FSW): the Frontiers Prevention Project (FPP) evaluation results*. BMC Public Health, 18; 10:497

Guy, R., Goller, J.L., Spelman, T., El-Hayek, C., Gold, J., Lim, M., Leslie, D., Tee, B.K., Roth, N., Anderson, J., Fairley, C.K., Kaldor, J. & Hellard, M. (2010). *Does the frequency of HIV and STI testing among men who have sex with men in primary care adhere with Australian guidelines?* Sex Transm Infect, 86(5): 371-6.

Halperin, D.T., de Moya, E.A., Pérez-Then, E., Pappas, G. & Garcia Calleja, J.M. (2009). *Understanding the HIV epidemic in the Dominican Republic: a prevention success story in the Caribbean?* J Acquir Immune Defic Sindr, 51 Suppl 1: S52-9.

Hamers, F.F. & Downs, A.M. (2004).*The changing face of the HIV epidemic in western Europe: what are the implications for public health policies?* Lancet ; 364(9428):83-94.

Heffelfinger, J.D., Swint, E.B., Berman, S.M. & Weinstock, H.S. (2007). *Trends in primary and secondary syphilis among men who have sex with men in the United States*. Am J Public Health, 97(6): 1076-83.

Istituto Nazionale di Statistica (ISTAT). (2005). Notification of Infectious Diseases in Italy-year 2002-Informazioni 2005; 2 : 9-15

Jin, F., Prestage, G.P., Zablotska, I., Rawstorne, P., Imrie, J., Kippax, S.C., Donovan, B., Templeton, D.J., Kaldor, J.M. & Grulich, A.E. (2009). *High incidence of syphilis in HIV-positive homosexual men: data from two community-based cohort studies.* Sex Health, 6(4): 281-4.

Katz, M.H., Schwarcz, S.K., Kellogg, T.A., Klausner, J.D., Dilley, J.W., Gibson, S. & McFarland, W. (2002). *Impact of highly active antiretroviral treatment on HIV seroincidence among men who have sex with men: San Francisco*. Am J Public Health, 92(3): 388-94.

Klausner, J.D., Wolf, W., Fischer-Ponce, L., Zolt, I. & Katz, M.H. (2000). *Tracing a syphilis outbreak through cyberspace.* JAMA, 284(4): 447-9.

Klavs, I., Bergant, N., Kastelic, Z. & Kustec, T. (2009). *Disproportionate and increasing burden of HIV infection among men who have sex with men in Slovenia: surveillance data for 1999-2008.* Euro Surveill, 14 (47):pii=19419

Kuklová, I., Velcevský, P. & Kojanová, M. (2011). *Syphilis among STD clinic patients in Prague in 2009.* Cent Eur J Public Health, 19(2): 84-90.

Kumta, S., Lurie, M., Weitzen, S., Jerajani, H., Gogate, A., Row-kavi, A., Anand, V., Makadon, H. & Mayer, K.H.(2010). *Bisexuality, sexual risk taking, and HIV prevalence among men who have sex with men accessing voluntary counseling and testing services in Mumbai, India.* J Acquir Immune Defic Syndr, 53(2):227-33.

Leber, A., MacPherson, P. & Lee, B.C. (2008). *Epidemiology of infectious syphilis in Ottawa. Recurring themes revisited.* Can J Public Health, 99(5):401-5..

Likatavicius, G., Klavs, I., Devaux, I., Alix,J. & Nardone, A.(2008). *An increase in newly diagnosed HIV cases reported among men who have sex with men in Europe, 2000-6: implications for a European public health strategy.* Sex Transm Infect, 84(6):499-505.

Marcus, U. & Hamouda, O. (2005). *Syphilis in Germany, 2004: diagnoses increasing, particularly in smaller cities and rural areas.* Euro Surveill, 10(7):E050728.3.

Matee, M.I., Magesa, P.M. & Lyamuya, E.F. (2006). *Seroprevalence of human immunodeficiency virus, hepatitis B and C viruses and syphilis infections among blood donors at the Muhimbili National Hospital in Dar es Salaam, Tanzania.* BMC Public Health, 6:21.

Mayer, K.H., O'Cleirigh, C., Skeer, M., Covahey, C., Leidolf, E., Vanderwarker, R. & Safren, S.A. (2010). *Which HIV- infected men who have sex with men in care are engaging in risky sex and acquiring sexually transmitted infections: findings from a Boston community health centre.* Sex Transm Infect, 86(1):66-70.

McCutchan, F.E. 2006. *Global epidemiology of HIV.* J Med Virol, 78 Suppl 1:S7-S12.

Merrigan, M., Azeez, A., Afolabi, B., Chabikuli, O.N., Onyekwena, O., Eluwa, G., Aiyenigba, B., Kawu, I., Ogungbemi, K. & Hamelmann, C. (2011). *HIV prevalence and risk behaviours among men having sex with men in Nigeria.* Sex Transm Infect, 87(1):65-70

Murray, J.M., McDonald, A.M. & Law, M.G. (2009). *Rapidly ageing HIV epidemic among men who have sex with men in Australia*. Sex Health, 6(1): 83- 86

Nagalo, M.B., Sanou, M., Bisseye, C., Kaboré, M.I., Nebie, Y.K., Kienou, K., Kiba, A., Dahourou, H., Ouattara, S., Zongo, J.D. & Simporé, J. (2011). *Seroprevalence of human immunodeficiency virus, hepatitis B and C viruses and syphilis among blood donors in Koudougou (Burkina Faso) in 2009*. Blood Transfus, 4(1):1-6. doi: 10.2450/2011.0112-10.

National Centre in HIV Epidemiology and Clinical Research. (2007). HIV/AIDS, Viral Hepatitis and Sexually Transmissible Infections in Australia Annual Surveillance Report 2007. National Centre in HIV Epidemiology and Clinical Research, The University of New South Wales. Sidney, NSW. Camberra: Australian Institute of Health and Welfare.

Pebody, R. (2010). *Internet proving a key way to reach men who have sex with men in Africa, Asia and beyond.* Available from http://aidsmap.com/page/1447379/

Peterman, T.A. & Furness, B.W. (2007). *The resurgence of syphilis among men who have sex with men*. Curr Opin Infect Dis, 20(1):54-9.

Prestage, G., Mao, L., Fogarty, A., Van de Ven, P., Kippax, S., Crawford, J., Rawstorne, P., Kaldor, J., Jin, F. & Grulich, A. (2005). *How has the sexual behaviour of gay men changed since the onset of AIDS: 1986-2003.* Aust NZJ Public Health, 29:530-35

Public Health Agency of Canada. (2007). *HIV/AIDS Epi Updates, November 2007, surveillance and Risk Assessment Division,* Centre for Infectious Disease Prevention and Control, Public Health Agency of Canada.

Sasse, A. & Defraye, A. (2009). *HIV infections and STI co-infections in men who have sex with men in Belgium: sustained increase in HIV diagnoses.* Euro Surveill, 14 (47) : pii=19420

Savage, E.J., Hughes, G., Ison, C. & Lowndes, C.M. European Surveillance of Sexually Transmitted Infections network. (2009). *Syphilis and gonorrhoea in men who have sex with men: a European overview.* Euro Surveill, 14 (47) pii: 19417.

Schneider, J.A., Lakshmi, V., Dandona,R., Kumar, G.A., Sudha, T. & Dandona, L. (2010). *Population-based seroprevalence of HSV-2 and syphilis in Andhra Pradesh state of India.* BMC Infect Dis, 10:59.

Simms, I., Fenton, K.A., Ashton, M., Turner, K.M., Crawley-Boevey, E.E., Gorton, R., Thomas, D.R., Lynch, A., Winter, A., Fisher, M.J., Lighton, L., Maguire, H.C. & Solomou, M. (2005). *The re-emergence of syphilis in the United Kingdom: the new epidemic phases.* Sex Transm Infect, 32 : 220-26

Smith, A.D., Tapsoba, P., Peshu, N., Sanders, E.J. & Jaffe, H.W. (2009). *Men who have sex with men and HIV/AIDS in sub-Saharan Africa*. Lancet, 374 (9687) : 416-223

Solomon, S.S., Srikrishnan, A.K., Sifakis, F., Mehta, S.H., Vasudevan, C.K., Balakrishnan, P., Mayer, K.H., Solomon, S. & Celentano, D.D. (2010). *The emerging HIV epidemic among men who have sex with men in Tamil Nadu, India: geographic diffusion and bisexual concurrency*. AIDS Behav, 14(5):1001-10..

Sullivan, P.S., Hamouda, O., Delpech, V., Geduld, J.E., Prejean, J., Semaille, C., Kaldor, J., Folch, C., Op de Coul, E., Marcus, U., Hughes, G., Archibald, C.P., Cazein, F., McDonald, A., Casabona, J., van Sighem,A. & Fenton, K.A. (2009). *Re-emergence of the HIV epidemic among men who have sex with men in North America, Western Europe, and Australia, 1996-2005.* Ann Epidemiol, 19(6):423-31

The White House Office of National AIDS Policy. National HIV/AIDS Strategy for the United States. Available from
http//www.whitehouse.gov/administration/eop/onap/.
Accessed on December 15, 2010

Velicko, I., Arneborn, M. & Blaxult, A. (2008). *Syphilis epidemiology in Sweden: re-emergence since 2000 primarily due to spread among men who have sex with men.* Euro Surveill, 13(50). pii: 19063

Wade, A.S., Kane, C.T., Diallo, P.A., Diop, A.K., Gueye, K., Mboup, S., Ndoye, I. & Lagarde, E. (2005). *HIV infection and sexually transmitted infections among men who have sex with men in Senegal*. AIDS, 19(18):2133-40.

Wallace, L., Winter, A. & Goldberg, D. (2005). *Increase in reported syphilis infections in Scotland in 2004.* Euro Surveill, 10 (30) : pii 2758

Ward, J.S., Guy, R.J., Akre, S.P., Middleton, M.G., Giele, C.M., Su, J.Y., Davis, C.A., Wand, H., Knox, J.B., Fagan, P.S., Donovan, B., Kaldor, J.M. & Russell, D.B. (2011). *Epidemiology of syphilis in Australia: moving toward elimination of infectious syphilis from remote Aboriginal and Torres Strait Islander communities?* Med J Aust, 194(10):525-9.

Watson-Jones, D., Oliff, M., Terris-Prestholt, F., Changalucha, J., Gumodoka, B., Mayaud, P., Semakafu, A.M., Kumaranayake, L., Gavyole, A., Mabey, D. & Hayes, R. (2005). *Antenatal syphilis screening in sub-Saharan Africa: lessons learned from Tanzania.* Trop Med Int Health, 10(9):934-43.

World Health Organization. (2001). *Global prevalence and incidence of selected curable sexually transmitted infections.* Available from
http// www.who.int/hiv/pub/sti/who_hiv_aids_2001.02.pdf

World Health Organization. (2007). *Sexually transmitted infections*. Fact sheet N°110. Available from http://www.who.int/mediacentre/factsheets/fs110/en/

Xu, J.J., Reilly, K.H., Lu, C.M., Ma, N., Zhang, M., Chu, Z.X., Wang, J.J., Yun, K. & Shang, H. (2011). *A cross sectional study of HIV and syphilis infections among male students who* *have sex with men (MSM) in northeast China: implications for implementing HIV screening and intervention programs.* BMC Public Health, 11: 287

Yang, Q., Boulos, D., Yan, P., Zhang, F., Remis, R.S., Schanzer, D. & Archibald, C.P. (2010). *Estimates of the number of prevalent and incident human immunodeficiency virus (HIV) infections in Canada, 2008.* Can J Public Health, 01(6):486-90.

Zetola, N.M. & Klausner, J.D. (2007). *Syphilis and HIV infection: an update.* Clin Infect Dis, 44(9):1222-8.

Zhang, M., Chu, Z., Wang, H., Xu, J., Lu, C. & Shang, H. (2011). *A Rapidly Increasing Incidence of HIV and Syphilis Among Men Who Have Sex with Men in a Major City of China.* AIDS Res Hum Retroviruses, Epub ahead of print.

24

Gun Violence in the United States: A Public Health Epidemic

Amy J. Thompson
The University of Toledo
USA

1. Introduction

Guns in America are not distributed evenly and once a firearm is purchased it is often retained by the owner for a substantial time due to its long-lasting design. It is estimated that about 34 million individuals own 195 million guns (Bureau of Alcohol, Tobacco, and Firearms, 2000). Estimates of private gun ownership are difficult to determine as there are few states that require firearm registration. There are two proxy methods that are commonly utilized to determine such estimates. Sales for the most widely sold gun magazine, *Guns and Ammo,* has been used to estimate gun ownership as well as the fraction of suicides in which a gun was involved (Duggan, 2001). Gun ownership seemed to peak in the early 1970's and since 1973 has declined. This decline would contradict the popular notion that gun ownership is escalating and that increased gun ownership is necessary to counteract the possibility of home invasion. While gun ownership has declined, handgun ownership has increased compared to shotguns and rifles (Blendon et al., 1996). Currently, an estimated 35% of households have at least one gun (Bureau of Alcohol, Tobacco, and Guns, 2000). Approximately, one in about four households contains a handgun (Smith, 2001).

According to a study by Cook & Ludwig (1997), the typical gun owner possesses multiple guns, is male, and is Caucasian. Gun ownership is highest among middle-aged, college-educated people in rural or small towns. A strong predictive factor of gun ownership is the presence of guns in an individual's home during childhood. Self-defense and recreation are the single most common reasons for purchasing a gun (Cook & Ludwig, 1997). A substantial minority of Americans (22%) report that they rarely leave home without a gun. On an average day, it is estimated that 1.1 million people carry a gun on their person outside of the workplace, while another 2.1 million store a firearm in their vehicle (Cook & Ludwig, 1997). According to Youth Risk Behavioral Surveillance System conducted by the Centers for Disease Control and Prevention (CDC) in 2009, six percent of high school students have carried a gun in the last 30 days.

The majority of guns purchased were obtained by their owners through federal firearm licensed dealers (FFL). However, a substantial number of guns (30% to 40%) are purchased through secondary sources that include gun shows, pawnbrokers, and newspaper classifieds (Cook, Molliconi & Cole, 1995). Under federal law, any gun purchased through a FFL requires a criminal background check. In most states, secondary sales do not require that the purchaser undergo a criminal background check. Secondary sales account for the majority of guns identified in crime traces (Cook, Braga, 2001). Another noteworthy source of gun acquisition is through theft. Each year more than 171,000 guns are stolen (Americana for Gun Safety, 2001)

The United States is not especially more violent than many other industrialized countries, yet the rate of firearm death is nine times greater than all other industrialized countries combined. In fact, the United States firearm death rate is three times higher than France, nearly five times higher than Israel, 12 times higher than Germany, 33 times higher than Great Britain, and 285 times greater than Japan (Krug, Powell & Dahlberg, 1998). Research suggests that it is ready access to guns that has drastically elevated the incidence and prevalence of firearm death and injury in the United States compared to other nations. The availability of guns has increased rates of violent crime across the United States. A study by Price, Thompson & Dake (2004) examined state variations in homicide suicide, and unintentional injury. They found in states with higher gun ownership there were also higher rates of firearm homicide, suicide, and unintentional injury. In other words, where there were more guns, there were more gun deaths. In a study by Webster, Gainer & Champion (1993), individuals who were more likely to carry guns were also more likely to be arrested, victims of violence, initiators of fights, and were willing to justify shooting someone.

Many people keep guns in readily available locations in their homes for personal protection. Research on police records involving home invasion data indicates that even in homes with guns, seldom are residents able to use a gun for self-defense. Despite this fact, many homeowners still perceive that owning a gunwill make the home safer. Howard, Webster, & Vernick (1999) conducted telephone interviews with a random sample of 4,138 individuals in urban areas. Twenty-nine percent of respondents believed that keeping a gun in the home would make them safer, 40% said less safe, 23% said that it depends, and 9% were unsure. It is quite possible, that the perceived benefits of having a gun for home protection are outweighed by the risks (Kellerman & Reay, 1986). Contrary to the idea that an individual will be more likely to use a gun to protect himself/herself or family members from an intruder a study by Kellerman & Reay (1986) indicated that for every case during a six year period in King County, Washington in which an individual used a firearm for a self-defense homicide there were 1.3 unintentional deaths, 4.6 criminal homicides, and 37 suicides involving guns. Other surveys have estimated that defensive gun use in the U.S. ranges from 61,000 to 23 million each year (Rand, 1994; Kleck & Gertz, 1995; Cook & Ludwig, 1997). Studies such as Kleck & Gertz's (1995) have estimated defensive gun use within the home at 800,000 per year but this figure has been criticized due to poor methodology (Hemenway, 1997). Data from California collected from 1990 through 1995 found that 17% of people murdered in a home were killed by a stranger. Of those killed in their homes, two-thirds knew their killer. When the shooting occurred in the shooters home, 63% of those killed were intimates or family members and 29% were acquaintances (California Department of Health Services, 1998). Thus, it would seem that the vast majority of firearm shootings occur among acquaintances rather than strangers.

1.1 Health and economic impact of guns

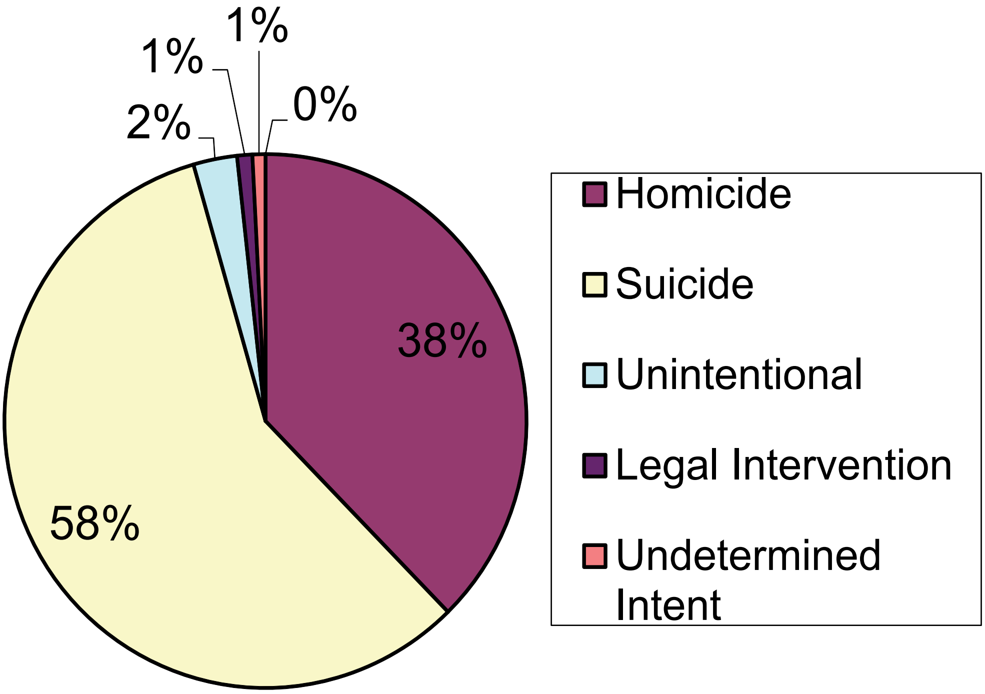
Each day in America more than 80 Americans die from firearm related injuries (CDC Wisqars , 2011).). Overall, guns are the second leading cause of fatal injury for Americans of all ages, second only to automobile accidents (CDC Wisqars , 2011). Suicide accounts for the majority (58%) of firearm fatalities, followed by homicide (39%) and unintentional injury (2%) (CDC Wisqars , 2011).

Males account for the vast majority (86%) of persons who die from gun trauma (CDC Wisqars , 2011). In African Americans, guns are the leading cause of death among the 15-24 age group (CDC Wisqars , 2011). In addition to the enormous human toll guns place on our society's health, the cost of treating these injuries impose a significant economic burden on our nations' taxpayers. In a study by Cook, Lawerence, Ludwig & Miller (1999) it was determined that the mean medical cost to treat a gunshot wound was about $17,000. Collectively in 2005, gunshot injuries in the United States cost $106,270,000 in medical expenditures. If you combine medical costs, with lost work time the overall expenditure is $27,673,032,000. Taxpayers end up paying high costs for medical care for victims of gunshot wounds due to the fact that this type of injury disproportionately impact those without private healthcare insurance. A study of gun-related hospitalizations in California revealed that 56% of patients hospitalized for firearm wounds relied on public health insurance, while 19% were totally uninsured (Vassar, Kizer, 1996). The most recent national data available suggests that the percentage of people who are injured by gun trauma who are uninsured has risen to 29% (Coben & Steiner, 2003).

1.2 Gun homicides

In the United States there were 12,179 homicides from guns in 2008 (CDC, 20011). Guns are often the weapon of choice by those committing a homicide. According to the Center for Disease Control and Prevention, 38 percent of firearm deaths are homicides (Figure 1). Homicide disproportionately affects young African American males. In fact, homicide is the leading cause of death for African American male youths. The gun homicide rate for African American men ages 18-29 is 133 per 100,000. This rate is about 25 times the rate for white males in that age group (Cook & Ludwig 2000). Gun homicide also disproportionately occurs in predominantly urban areas. If a gun homicide does occur in a rural area, it is more likely to involve a white male (Cook et al., 2001).

Although guns are more likely to result in completed suicides than homicides, it is criminal homicides that often cause the greatest public concern. Although many individuals purchase a gun for personal protection against possible homicide or crime there is a significant relationship between those who purchase a firearm and the risk of their own homicide (Cummings, Koepsell, Grossman, Savarino,& Thompson, 1997). Cummings and colleagues, (1997) found that homicide victims, when compared to controls, were more likely to have either purchased a handgun or have a family history of a handgun purchase. The connection between the family history of a handgun purchase and homicide is not surprising considering that the majority of homicides occur either between two people who are family members or who are acquaintances (Bureau of Justice Statistics, 2001). Thus, contrary to popular notion, the person who is most likely to kill you is someone you know as compared to a stranger. Weapon selection is often linked to the likelihood of a homicide occurring. In many attacks, the assailant often lacks a clear or sustained mind or intent to kill (Cook et al., 2001). Whether the victim lives or dies often depends on the lethality of the weapon chosen. Unfortunately, for many victims, guns are the weapon of choice by their attacker. Compared to many other possible weapons, a gun can be fired from a significant distance and often does not give the victim the opportunity to overcome the assailant. Guns can also inflict more injury with relative ease when compared to other weapons (Cook et al., 2001)



Data Source: CDC, 2011

Fig. 1. 2008 Gun Death by Intent in the United States

2. Firearm suicides

Fir suicides have dramatically increased since the 1950's. According to the Centers for Disease Control and Prevention, (2011) in the United States, more people kill themselves with guns than with all other methods combined (Figure 2). In 2008, there were approximately 36,035 suicide deaths among Americans and 50% of these deaths were caused by guns (CDC Wisqars , 2011) . Community suicide rates are greatly influenced by gun availability (Markush & Bartolucci, 1984; Miller & Hemenway, 1999). Brent and colleagues (1994) provided data to suggest that access to guns is a critically important risk factor for adolescent suicide. In their analysis, they found that guns were twice as likely to be found in the home of suicide victims than in the home of suicide attempters. In other words, those who actually completed suicide were more likely to use a gun. Both suicide rates and unintentional gun injuries are increasing due to availability and access to guns (Miller, Azrael & Hemenway, 2001). Contrary to popular belief, more people die from gun suicide than from gun homicide. In 2011, there were 1.5 times as many suicides as homicides (CDC Wisqars , 2011).

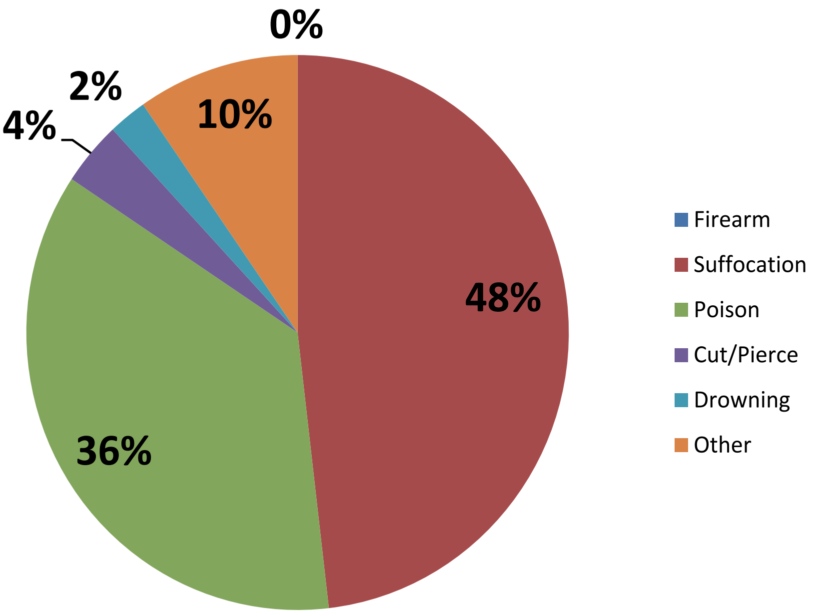
Guns account for more completed suicides than any other means for men, women, and children. This should not be surprising considering due to the lethality of guns compared to other methods. Expert opinion and related evidence support the idea that both intent and instrumentality matter, individuals who commit suicide often do so when confronting a severe but temporary crisis (Seiden, 1977). For example, a study conducted by Chapdelaine, Samson & Kimberly (1991) found that 92% of gun suicide attempts resulted in death compared to 78%

of attempts using carbon dioxide or hanging, 67% using drowning, and 23% of intentional drug overdoses. A further study by the Centers for Disease Control and Prevention (1997) found that for every uncompleted gun suicide there were five completed gun suicides.

Impulsiveness appears to play an important role in suicide, especially among youth suicides. It is not uncommon for adolescents to have passing suicidal impulses and the availability of a gun increases the likelihood suicide will be completed. Research has determined that in many cases of suicide attempts if the individual survives the initial attempt then he/she did not have a true sustained desire to die (Blumentahl & Kupter, 1990).

Access to a firearm in these situations may prematurely end someone's life who never intended to die. For example, in one study of 18 men who survived self-inflicted intentional gunshot wounds to the face, subsequent suicide attempts were uncommon (Miller, 2001). Another study by Chapdelaine and colleagues (1991) examined self-inflicted gunshot wounds and the impulsiveness that may occur in those contemplating suicide. Out of the 30 suicide attempters observed, none had written a suicide note, and more had reported suicidal thought for less than 24 hours. Subsequent follow up of these individuals revealed that after two years, none of the 30 people attempted suicide again.

Kellerman and colleagues, (1992) found that people living in a household with a gun are almost five times more likely to die by suicide than people who live in gun-free homes. Miller (2001) also confirmed the risk of household gun ownership and suicide using state level data from 1988-1998. In this particular study, there was a strong statistical relationship found between household gun ownership and suicide. This relationship could be seen in both genders and for every age group. These findings again suggest that having a firearm available for ready access may increase the chances of injury whether it be through homicide or suicide.



Data Source: CDC, 2011

Fig. 2. Method of Suicide in the United States 2008

While access to guns may contribute to suicide completion one of the largest groups impacted by firearm suicide is adolescents. In 2008, 748 young people ages 19 and younger used a gun to commit suicide (CDC Wisquar, 2011). Forty percent of young people who commit suicide use a gun (CDC Wisquar, 2011). As previously discussed, impulsiveness may play a key role in suicide death but this may be particularly true for adolescents who may contemplate suicide. In a study by Shah, Hoffman, Wake and Marine (2000) of 36 adolescents who committed suicide, 67% did so with a gun obtained from their own home. Therefore, a recommendation to decrease the risk of suicide for adults and adolescents would be to limit household access to guns (Shah et al., 2000).

3. Unintentional gun injury

In the year 2008, there were 592 unintentional gun deaths in the United States (CDC Wisqars, 2011). White males between the ages of 15-19 years account for the majority of all unintentional firearm deaths. The majority of urban unintentional firearm injuries occur with a handgun while a large proportion of rural injuries are inflicted with long-guns (General Accounting Office, 1991; Sadowski & Munoz, 1996). Overall, unintentional gun deaths account for a relatively small portion of all gun deaths (approximately 2%). However unintentional firearm injuries impact a substantial number of young Americans (National Center for Injury Prevention and Control, 2002). While difficult to ascertain due to lack of adequate data collection, nonfatal firearm injury is significantly higher than individual unintentional gun death, homicide, and suicide rates. For example, in 2009 there were an estimated 18,610 people in the United States treated in hospital emergency rooms for nonfatal, unintentionally inflicted gunshot wounds (CDC Wisquars, 2011).

The rate of unintentional gun deaths has declined steadily throughout the 20th century (Ikeda, Gorwitz, James et al., 1997). This decline may be significantly impacted through the implementation of gun safety features and safe storage practices (Frattarolli, Webster & Teret, 2002). Safe storage of guns in particular may be beneficial in the prevention of child unintentional gun injuries (Teret & Baker, 1995). Many children are killed by guns when handling guns in the home. Knight-Bohnoff & Harris (1998) surveyed 80 parents of children in a daycare setting. They found that 48% of parents kept at least one gun in the home. Twenty-six percent of those gun owners reported that the gun was loaded at all times, and 18% reported that a gun was kept within the reach of a child. In another study by Nelson, Grant-Worley, Powell et al., (1996), 10% of adults who lived in households with guns stated that it was always or sometimes stored loaded and unlocked. This access accounts for the significant number of unintentional gun injuries among youth despite the steady declines. Some may argue that the reduction in unintentional gun injuries and deaths over the last century may have been a result of gun safety education that taught children not to touch a gun and how to be safe around guns. Yet, one study by Hardy, Armstrong, Martin & Strawn (1996) found that a gun safety education program did not reduce youth's tendencies and curiosities to touch a gun when they were not supervised.

4. Safety mechanisms to reduce unintentional firearm injury

There are a number of technologies that exist for making guns safer yet these mechanisms are not commonly implemented by gun manufactures or purchased by consumers. Several of these options are designed to limit access to "high-risk" users such as children, youths, and home burglars (Frattaroli et al., 2002). One type of device designed to reduce unintentional injuries is the chamber indicator. In many handguns the handle or the grip of the gun containing the ammunition is in either a magazine or clip. In order to fire the weapon the shooter must move the slide of the firearm to bring the cartridge into firing position. During this process it is not obvious whether there is a cartridge already in the chamber because the chamber is not viewable by the user. This danger can be easily corrected by equipping all pistols and revolvers with a device that indicates whether ammunition is in the chamber (Frattaroli et al., 2002). Currently, few guns are available with such a device (Vernick, Meisel, Teret et al, 1999).

Another useful safety device is a magazine safety or disconnect. When an ammunition magazine is removed from the gun a live bullet can still remain in the chamber. Even though the magazine is not connected to the gun the weapon can still be fired with the remaining bullet. A magazine disconnect disables the gun from firing when the magazine is removed from the gun. Although patents for this device have been established for nearly a century few handguns are equipped with this technology (Frattaroli et al., 2002; Vernick et al., 1999).

Of the various techniques available to modify a gun, personalization of guns has the greatest potential to reduce the largest number of firearm deaths. Personalized guns can only be fired by the designated operator. There are a number of ways that a gun can be "personalized". These methods may include manufacturing guns with a built in combination lock or a removable trigger lock, requiring a personal identification number, or even requiring the owners fingerprints to fire the weapon (Frattaroli et al., 2002;Vernick et al., 1999). Thus, children who found a gun in their home would be unable to fire the weapon causing unintentional injury, adolescents would be prevented from accessing it to commit suicide, and criminals would be unable to use a stolen gun to commit a crime such as a homicide. Each year many unintentional gun injuries result from handling loaded guns. These safety devices have the potential to reduce a significant number of injuries, yet for the most part are not being implemented by guns manufacturers.

5. Gun control policy debate

Gun control is a term that encompasses many dimensions and issues. One of the most significant debates involving gun control is how some restrictions involving private ownership may violate the Second Amendment of the Constitution (Blendon et al., 1996). Other areas of debate regarding guns include: gun sales to those deemed dangerous, gun registry, and banning of specific types of guns such as machine guns, assault weapons, handguns, and manufacturing of cheaply made guns coined "Saturday Night Specials" that are easily concealed and disposed of (Blendon, et al., 1996).

Citizens differ widely in their positions on such debates. Much of the public favors gun control measures that do not impose major restrictions on gun owners. The public significantly opposes gun sales to convicted felons or individuals under the age of 18 years of age. A vast majority of individuals also support the Brady Act a law that imposes background checks and a five-day waiting period. Finally, there is also significant public support for registration of handguns and limiting gun purchases to 1 per month (Blendon et al., 1996).

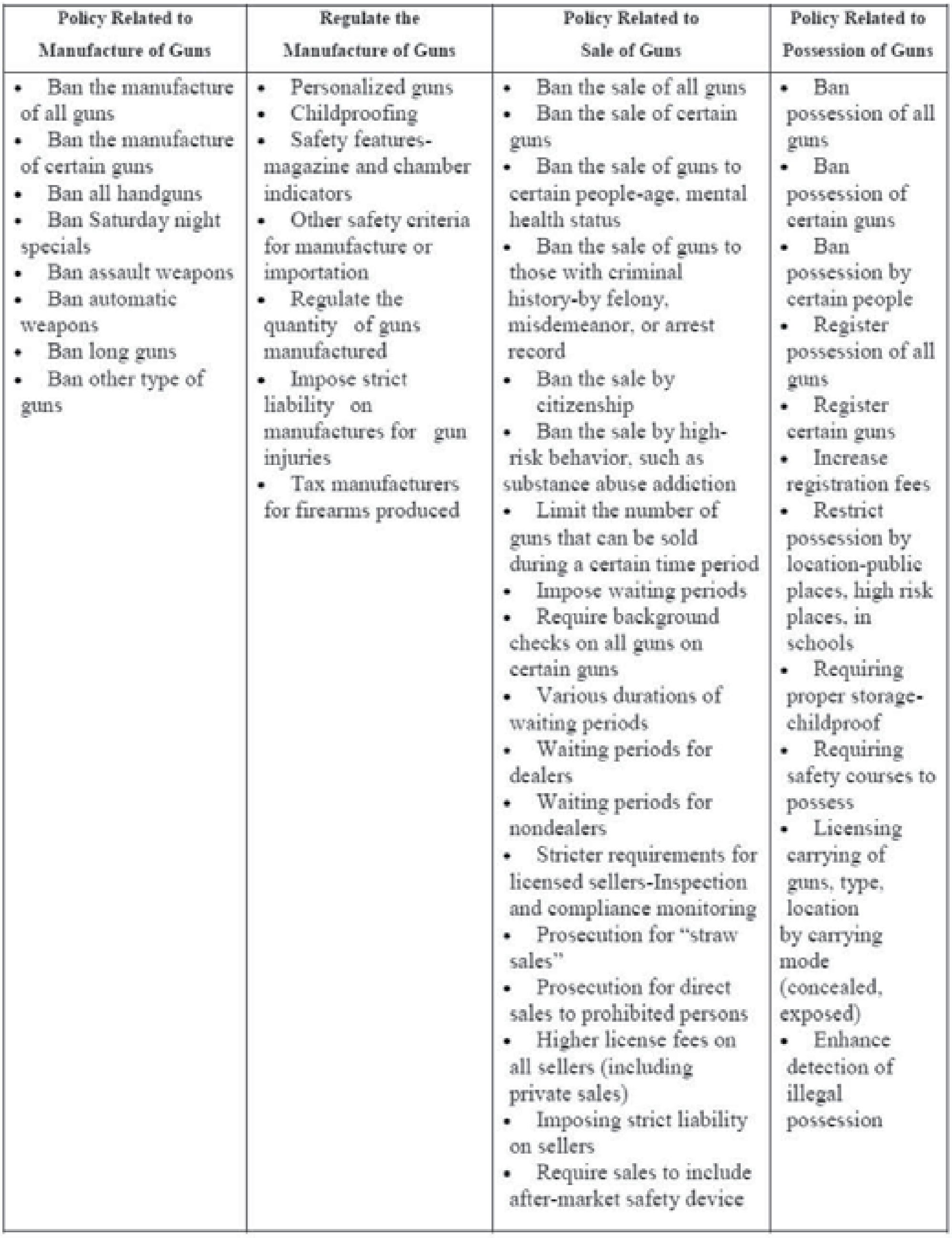
In a study by Teret and colleagues (1998), 2400 adults were surveyed concerning their views towards firearm regulation. They found that a majority of respondents favored safety standards for new handguns. These standards included childproofing (favored by 88%), personalization (devices that permit firing only by an authorized person; 71%), magazine safeties (82%), and load-chamber indicators (73%). There was also a strong public support for policies prohibiting convicted felons from obtaining guns (85-95%) (Teret at al., 1998). There was also significant support for policies designed to reduce the illegal sale of guns, including tamper resistant serial numbers (90%), limit of hand gun sales to one per month (81%), and mandatory registration of handguns (82%) (Teret et al., 1998). This study is significant because it showed strong public support, even among gun owners, for new strategies to regulate guns and make them safer. With such overwhelming public support it seems questionable why stricter firearm control policies are not being enacted at the federal and state level.

With so many types of ammunition and guns available it is difficult to establish gun laws that encompass all of these products. Consumer trends in gun purchasing are more clustered around handguns than ever. This type of weapon is used much more often in homicides, particularly fatal ones (Johnson, Coyne-Beasley & Runyan, 2004). Handguns are both very powerful and easy to conceal. Gun manufacturers know there is a demand in this area. They are producing smaller more easily concealed handguns. In addition to this deadly design, semiautomatic handguns are being produced to rapidly reload and fire. Other products deemed "point and shoot accessories" such as laser sightings have been developed to ensure that both the target is not missed and that a clear recognition of a target is not necessary for accuracy (Wintemute, 1996).

5.1 State and federal gun laws

It is estimated that there are at least 300 state and federal firearm laws (Vernick, Hepburn, 2002). There have been a number of categories of gun policies that have been proposed in the literature. Some researchers suggest that restrictions should focus on the manufacture, sale, and possession of guns (Baker, Teret & Dietz, 1980). Zimring & Hawkins (1987) proposed gun policies that focused on restricting guns in certain places; severity of penalties, licensing, registration, and ban considerations. Kellerman (1991) has suggested applying a public health approach to firearm injury prevention. Christoffel (1991) has proposed a number of safety polices such as child access laws for preventing gun injuries in children. There are a number of firearm control policies that have been adopted at the federal and state level. Figure 3 adapted from Teret & Wintemute (1993) provides a list of possible categories and sub categories of firearm policies .

One of the first federal laws enacted to impact the sale of guns was the National Guns Act of 1934. This law was enacted during the prohibition era and was inspired by the attempted assassination of President-elect Franklin Roosevelt (Brady Campaign, 2002). It placed a tax on manufacturers, and restricted the sale of sawed off rifles, shotguns, machineguns, and silencers. In 1938 the Federal Guns Act was passed which required annual licenses for manufacturers, dealers and importers of guns and ammunitions. Nearly 30 years later, with the assassination of Martin Luther King Jr. and Robert Kennedy in combination with rising crime rates, the Gun Control Act of 1968 was passed. Under this law, the sale of guns by dealers were banned to certain categories of people (for example, minors, convicted felons, mentally ill people, drug addicts, fugitives, undocumented immigrants, and anyone who was dishonorably discharged from the military). This law also required mandatory placement of serial numbers on all guns (Brady Campaign, 2002).



Policy Related to Manufacture of Guns	Regulate the Manufacture of Guns	Policy Related to Sale of Guns	Policy Related to Possession of Guns
• Ban the manufacture of all guns • Ban the manufacture of certain guns • Ban all handguns • Ban Saturday night specials • Ban assault weapons • Ban automatic weapons • Ban long guns • Ban other type of guns	• Personalized guns • Childproofing • Safety features-magazine and chamber indicators • Other safety criteria for manufacture or importation • Regulate the quantity of guns manufactured • Impose strict liability on manufactures for gun injuries • Tax manufacturers for firearms produced	• Ban the sale of all guns • Ban the sale of certain guns • Ban the sale of guns to certain people-age, mental health status • Ban the sale of guns to those with criminal history-by felony, misdemeanor, or arrest record • Ban the sale by citizenship • Ban the sale by high-risk behavior, such as substance abuse addiction • Limit the number of guns that can be sold during a certain time period • Impose waiting periods • Require background checks on all guns on certain guns • Various durations of waiting periods • Waiting periods for dealers • Waiting periods for nondealers • Stricter requirements for licensed sellers-Inspection and compliance monitoring • Prosecution for "straw sales" • Prosecution for direct sales to prohibited persons • Higher license fees on all sellers (including private sales) • Imposing strict liability on sellers • Require sales to include after-market safety device	• Ban possession of all guns • Ban possession of certain guns • Ban possession by certain people • Register possession of all guns • Register certain guns • Increase registration fees • Restrict possession by location-public places, high risk places, in schools • Requiring proper storage-childproof • Requiring safety courses to possess • Licensing carrying of guns, type, location by carrying mode (concealed, exposed) • Enhance detection of illegal possession

Fig. 3. Possible Gun Polices

Twenty-five years later, a presidential assassination attempt sparked the enactment of a federal gun control law. In 1981, President Ronald Reagan was shot by John Hinckley Jr. During the assassination attempt Press Secretary James Brady was severely wounded. In 1993, the Brady Law was enacted named in honor of James Brady. Prior to the enactment of this law many states did not require a criminal background check of those who desired to purchase a firearm. The Brady Law changed this by requiring all federally licensed firearm dealers (FFL's) to complete a background check with law enforcement prior to selling a firearm. People who have guns through secondary gun sales however, are not required to undergo criminal background checks. Under this law, guns are also prohibited from being sold to those who are illegal aliens, fugitives, those adjudicated as mentally defective, those who have been issued a restraining order, and those convicted of a domestic abuse misdemeanor. Handgun sales are also not permitted to those under the age of 21 and long guns cannot be purchased by those under the age of 18 (Brady Campaign, 2002).

In addition to federal law, each state can enact its own gun laws. However, many states have preemption laws which do not permit the adoption of local ordinances. The states that do not have preemption include Alaska, California, Colorado, Florida, Hawaii, Illinois, Kansas, Maryland, Massachusetts, Montana, Nebraska, New Hampshire, New York, and Ohio (Guns Law Center, 2002). The effectiveness of many of these laws can only be speculated as the majority of them have not been rigorously evaluated. In 1978, a report was issued to Congress by the Controller General regarding the effect of gun laws on violent crime. The report suggested:

> *"Though much has been written on the subject of controlling guns. Only a few good empirical studies evaluate the impact of gun control laws on violent crime….Of these studies, some attempted to show that gun control reduces homicide and some attempted to quantify the effect of different types of gun control laws. Several of these studies conclude that gun control laws, through limiting firearm availability, result in decreased murder. None of the studies demonstrate a decrease in overall level of violence attributable to gun control."* (Teret & Wintemute, 1993 p.101)

Subsequently since that report was issued, the American Medical Association (AMA) Council on Scientific Affairs issued a report of its own regarding the effectiveness of gun-control measures. This report suggested that the problem of guns is complex and is one of the more controversial issues facing the public health profession in recent years and that more information needs to be gathered on the effectiveness of firearm control measures (AMA, 1989). This may be due to the large gaps in the evaluation of gun laws, variations of laws from state to state, and degrees of enforcement. Very few laws regulate the manufacture of guns, thus policy evaluation in that area is limited (Teret, Wintemute, 1993). Most research that has been done on the effectiveness of gun laws have focused on regulating the sale and possession of guns, and sentence enhancements when crimes are committed with a gun.

5.2 Empirical research on the effectiveness of gun policy

At the federal level there have been few studies that have examined the effectiveness of gun policy. Ludwig & Cook (2000) assessed homicide and suicide rates associated with implementation of the Brady Handgun Violence Prevention Act. The Brady Handgun Violence Prevention Act was implemented in February 1994 and required those 21 years of age and older to undergo a background check and five day waiting period for those seeking to purchase a handgun from a federally licensed dealer. When this law went into effect, 32 states were required to conduct background checks and to undergo the five-day waiting period. From 1994-1997 the impact of the Brady Law was followed and it was found that changes of homicide and suicide for those states who implemented the Brady Law were not significantly different from those states that were not required to implement the law. The Brady Law did seem to have been associated with a reduction in suicides by those over the age of 55. This may be due to the waiting period requirement. Because the Brady Law was changed with the implementation of the instant background check system and was only in place for a few years it may be difficult to ascertain any potential benefits this federal law would have made had the waiting period portion of this law remained intact.

At the state level, restrictions of the sale of guns may prove to have some merit in reducing crime and injury. Webster, Vernick & Hepburn (2001) examined the relationship between licensing, registration, and other gun sale laws in regards to the source of state crime guns. They found that states with registration and licensing systems were less likely to have guns sold within the state that were recovered in crimes. Weil & Knox (1996) examined the effects of limiting handgun purchase though one handgun per month on the interstate transfer of guns, and the effects of licensing and registration. They also reported a decrease in state level crime. In 2002, Webster, Vernick & Hepburn assessed the effect of the 1988 law that the Maryland legislature passed which limited the sale of certain types of small concealable handguns known as "Saturday Night Specials". This study found a delayed or gradual effect of the ban that produced estimates indicating that firearm homicide rates were 6.8-11.5% lower than would have been expected without the ban.

A number of studies have researched the outcomes of limiting the possession or access of guns. Cummings, Grossman, Rivara & Koepsell (1997) found a relationship between child access prevention (CAP) laws that make guns less accessible to children by holding adults criminally responsible for youth access. In those states where CAP laws have been in existence for at least one year, unintentional gun deaths were reduced by 25%. A follow-up study by Webster & Starnes (2000) utilized a pooled time series model to determine validity of the study by Cummings and colleagues. Webster and colleagues (2000) found a reduction in unintentional injury death among those 15 years and younger, but the reduction was a smaller decline (17% vs 25%) than found by Cummings and colleagues (1997).

Some states have banned the possession of certain types of guns as well as having placed restrictions on how guns can be carried. In 1976, the District of Columbia passed the Guns Control Regulations Act, which prohibited the purchase, sale, transfer, and possession of handguns by civilians unless they previously had owned a handgun and registered it. Loftin, McDowell, Wiersma & Cottey (1991) evaluated the impact of this law and that passage of the law coincided with a decline of about 25% in the rates of homicide and suicide by gun. As a measure of validity, surrounding metropolitan areas that did not have the restrictions imposed were studied during the same period of time. There were no similar reductions in homicide and suicide by guns found and thus the authors concluded the law prevented an average of 47 deaths per year after the law was implemented. This effect persisted through 1987, which may have resulted in more than 500 lives being saved (Teret et al., 1993). Also, during the early 1970's Massachusetts adopted a law known as the Bartley-Fox Amendment. This law strengthened handgun-licensing protocol, and established a one-year mandatory prison sentence for carrying a gun that was not properly licensed. Under this law, all of those convicted must serve their entire sentence and were not eligible for early parole or probation. Beha (1977) evaluated the effectiveness of this law and found that a publicized threat of mandatory sentencing for carrying an unlicensed firearm was effective in deterring crime. Pierce and Bowers (1981) in a follow up study of the same law also found that the law reduced the number of gun assaults however during this same time the number of non-firearm assaults increased.

Several studies have examined the impact of concealed carry weapons laws (CCW) (Lott & Mustard, 1997; Ludwig, 1998; McDowall, Loftin, Wiersma, 1995; Ayres & Donohue, 2002). Currently only 1 states(Illinois) entirely prohibit the carrying of a concealed weapon. "Shall issue" laws require law enforcement agencies to issue concealed weapon permits to any applicant who meets specific criteria such as minimum age, free from criminal convictions, or having a diagnosed mental illness. "May issue" laws are more restrictive and require law enforcement agencies to issue permits to individuals who qualify under limited criteria such as individuals who transport large sums of money or who work as a private investigators.

While one of the common arguments for CCW laws is to promote public safety through individual protection with a firearm, the opposite has occurred in states that have initiated such legislation. For example, a study by the Violence Policy Center (2002), revealed in Texas (one of the few states that allows the tracking of permit holders) "law abiding" permit holders were at a rate of two and a half crimes per day since the law went into effect. From 1996 to 2000, Texas concealed handgun license holders were arrested for weapon-related offenses at a rate 81 percent higher than that of the general population of Texans ages 21 and older.

Crime and violence may be easier to track in Texas but other states have found similar results after passage of CCW laws. In a study by McDowall, Loftin & Wiersma (1995) five cities were examined in three different states that had existing CCW laws. They found that shall-issue laws were associated with significant increases in gun homicides in three of the five cities. They also found that Florida's shall-issue law was associated with an increase in homicides for the entire state.

Unfortunately, there is currently little research on the impact of "shall-issue" laws and crime. Perhaps one of the most cited pieces of literature utilized by proponents of CCW laws is Lott and Mustard's (1997) study on "Crime, Deterrence, and Right to Carry Concealed Handguns". The authors of this study suggested that concealed weapons "shall issue" laws may be an effective method in reducing crime. Moreover, these researchers proposed that more guns equate to less crime. Lott and Mustards (1997) research however has been the focus of numerous empirical studies that questioned their findings (Black and Nagin, 1998; Webster, Vernick, Ludwick, Lester 1997, Donahue and Ayres, 2003). Lott and Mustards (1997) findings are not supported by other research because of serious flaws in the study design, which biases the results toward finding crime-reducing effects. Some of the other criticisms of their study include measurement error, failure to control for variables, and omitting potentially important variables.

A study by law professors John Donohue of Stanford and Ian Ayers of Yale (2003) also contradicts the findings of Lott and Mustard (1997). One of the key findings of this study was that jurisdictions that were analyzed had an increase in crime in the "shall issue" permit states. Robbery is committed in public more than any other crime and should be the crime most likely to decline if the deterrence of CCW is effective. Yet in the states that have not adopted shall-issue laws there is a much lower rate of robbery than states that allow the carrying of concealed handguns

Ayres and colleagues (2003) also hypothesize that even if no one securing a concealed carry permit ever used it to commit a crime, there are still a number of avenues that crime can increase. First, even if the adoption of a shall issue-law increased the riskiness of criminal activity, it might also increase the number of criminals who would illegally carry weapons themselves, and might also increase the speed at which a criminal decides to shoot or disable a victim (as the presence of an armed victim decreases hesitation once a criminal engagement has been launched). Therefore, the number of murders and aggravated assaults might rise if criminals responded to shall-issue laws by packing more guns, carrying higher-powered weapons, and discharging them more quickly when threatened. Second, the massive theft of guns each year means that anything that increases the number of guns in America will likely increase the flow of guns into the hands of criminals, who may in turn use them to commit crime.

Proponents of CCW contend that there is an inherent need for firearm protection in public places. The typical gun permit holder however is a middle- aged white male, which is a group at relatively low risk of violent criminal victimization with or without gun ownership. Therefore it is not clear whether substantial benefits are likely to occur by further arming this group (Donahue et al., 2003). There is the perception that carrying a concealed weapon can help one to protect oneself from an unknown offender and that one could save others as well in a dangerous situation such as a robbery or a hijacking. The reality is that by having a gun in an individual's hands who is not trained to respond to such a situation could cause more harm than good. Citizens are often not adequately trained to handle guns in stressful situations. According to the FBI, in 1999, 5 out of 41 law enforcement officers killed in the line of duty were killed by an adversary with the officer's own service weapon (FBI, 1999). If this can occur to a police officer it is perhaps even more likely to happen to a lay citizen.

Other researchers have found that concealed carry laws actually increase homicide and crime rates. Ludwig (1998) analyzed state level data on concealed carry laws and found that shall issue laws have resulted in an increase in adult homicide rates. McDowall, Loftin, and Wiersema (1995) conducted a multi-state comparison on the effects of concealed weapon laws (Florida, Mississippi, and Oregon). Across the major counties studied in the three states, gun homicides increased in the aftermath of "shall issue" laws. In contrast, during the same period, homicide without weapons remained consistent. These findings were consistent even when other potential confounding variables were controlled for such as population change, variations in national homicide rates, and the enactment of other laws within the state. It was found that after the passage of the "shall issue" laws in those states that the average number of homicides in the major urban areas within those states increased by 4.5 persons per 100,000 (McDowall et al., 1995).

Studies conducted outside of the United States may also provide information on the effects of concealed weapons laws. Villaveces, Cummings & Espitatia et al., (2000) assessed the effect of a ban on carrying guns on homicide rates in two Colombian cities. In 1983-1993 there was a 366% increase in the homicide rate in Columbia. To combat this increase, the carrying of guns was banned on weekends after paydays, on holidays, and on election days.

Police established checkpoints to monitor compliance. During periods when the ban on carrying concealed guns was enforced homicide rates were lower than expected with an adjusted odds ratio of 0.86.

As previously stated, the United States has the highest gun death rates in the industrialized world. One of the countries that often is compared to the United States for gun homicide and suicide is Canada. Canada has a significantly lower number of gun deaths compared to its geographic neighbor, the United States. One speculative cause of this lowered gun death rate is the stricter gun control laws that are in effect in Canada. Le Leenaars and Lester (1993) examined the effects on homicide of Canada's Criminal Law Amendment Act of 1977, which has been in effect since 1978. This law requires acquisition certificates for all guns, restricts access to certain types of guns such as handguns, reduces the availability to certain types of individuals such as felons, sets up procedures for handling and storing guns, requires permits for sellers of guns, and increases sentences for firearm offenses. Leenaars and colleagues (1993) examined Canadian homicide data from 1969 through 1985. Whereas the homicide rate during this time period did not significantly change after the passage of Canada's Criminal Law Amendment Act, the number of homicides committed with a firearm decreased from 38 percent in 1977 to 33 percent in 1985. This reduction in firearm homicides may be encouraging but there was some evidence to suggests that because the overall homicide rate in Canada did not decline that individuals intent on murders may have switched to other means.

Australia is another country in which there has been some tracking of the effectiveness of gun policy. After a lone gunman opened fire in Port Arthur and killed 35 persons, the Australian state and federal government implemented a broad plan for gun control. Within 12 months of the massacre, certain types of guns were prohibited (semiautomatic and self-loading) and new licensing, registration, safe storage, and safety training requirements were implemented. These new policies coincided with a massive gun buy-back of those weapons that were now banned. The premise behind such a buy-back program was that by reducing the number of lethal guns that gun crime, mortality, and morbidity would be reduced. After the implementation of such a buy-back program in conjunction with the other gun polices previously listed Australia did have a reduction in the number of homicides by 10 percent (Reuter & Mouzos, 2003). Overall, the impact of all of these combined measures did not decrease the total number of suicides but did reduce the number of suicides committed with a gun. Violent crime also appears to have been impacted by these new policies. Although data suggests that the rate of attempted murder remained stable (29 percent in 1993 and 31 percent in 2000) the rate of guns used in assaults become more rare (Reuter et al., 2003). For example, in 2000 there were 141, 24 assaults, of which 786 involved the use of a gun. This translates into barely 0.5 percent of assaults being committed with a gun.

While it appears that there has been some reduction in gun homicide and violent crime in Australia after the enactment of these policies it is worth noting that research in the United States regarding gun buyback programs has been less than promising (Rosenfield, 1996). The failure of gun buy backs programs has often been attributed to the small scale of the interventions. Reuter et al., (2003) gives the example that if a gun buy-back program had a $100,000 budget, and for each gun turned in $50 (a typical price) was distributed, the total number of guns collected would be 2,000. This would equate to less than one percent of the guns in a city with a population of 250,000 (Kleck, 1996). The apparent difference between the gun buy-back programs in Australia and the United States is both the dollar amount given for the gun turned in and the scale of the campaign. In Australia the average dollar incentive awarded per gun turned in was $359. Since this program has been in effect nearly 4.25 million guns were turned in decreasing gun ownership by 20 percent (Reuter et al., 2003).

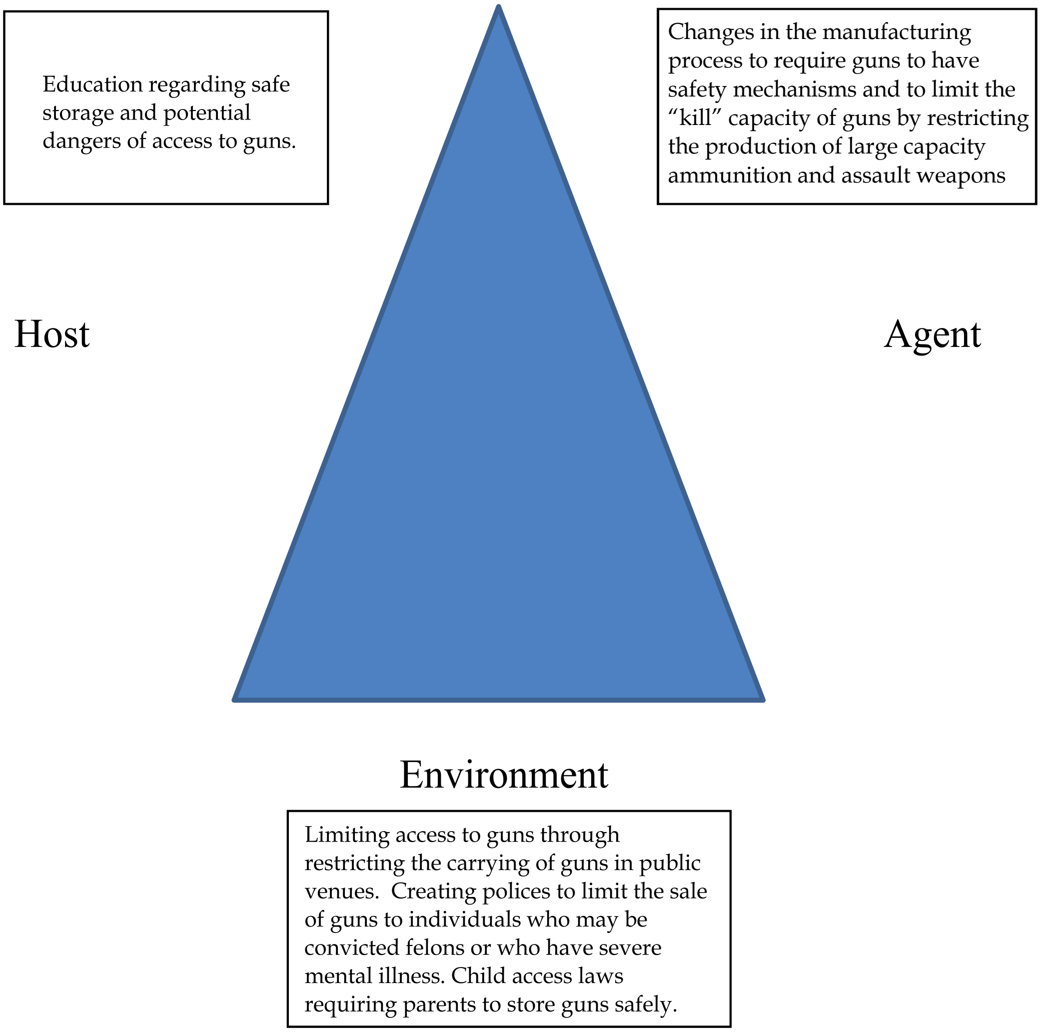
While there has been various studies published regarding the effectiveness of various gun laws the Task Force on Community and Preventive Service (2005) has completed one of the most comprehensive and complete examination on the effects of the major gun laws. This group specifically examined bans on specific types of restrictions on gun acquisition, waiting periods for firearm acquisition, firearm registration and licensing of gun owners, CCW laws, child access laws and zero tolerance laws in schools. After careful review of the research conducted the committee issued a comprehensive report indicating there was insufficient evidence that any of the aforementioned gun laws were indeed effective. This conclusion was based on several reasons that include: too few studies, unreliable data collection methods that do not control for confounding variables, inappropriate analysis and inconsistent results

5.3 Alternatives to the enactment of gun policy

While the enactment of several gun laws in the United States has been successful many gun laws are too narrow to dramatically impact public health. Koop and Lundberg (1992) suggest that both gun control and violence prevention needs to encompass many different strategies. They make three recommendations: support additional research on the causes, prevention, and cures of gun violence; increase the education of all Americans about violence, and demand legislation to reverse the upward trend of injuries and deaths. Further, they recommended that a legislative approach similar to that of injury control of motor vehicles be mandated for guns. This approach encompasses modification of host, agent and environment (Figure 4)

Teret and colleagues (1993) proposed that for optimal reduction in firearm deaths a redirection of policy away from attempting to regulate the use and possession of guns towards regulating their manufacture, sale and marketing. This may be particularly successful at the state level. They argued in doing so, successful injury control could be achieved. In addition to policy implementation there may be additional strategies that may help to reduce the number of people who are killed or injured from guns. Patient counseling by the medical community may provide alternatives when firearm policy or regulation is not feasible or in addition to adoption of gun policy.

Multisite means reduction and safe storage education may be one of the most effective ways to reduce firearm injury. When patients are counseled by members of the medical or mental health community behavioral change is often observed. Such success can be seen with patients regarding smoking cessation, seatbelt use, and diet. This often is due to patients deferring to the advice of a perceived expert such as a physician, physician assistant, nurse, psychiatrist or counselor. Several studies have examined the influence and the extent of gun counseling behaviors by professionals. A study of adults' perceptions of physicians' advice with regard to not keeping guns in the home found that 47% claimed that they would follow the advice, and an additional 37% would think over the advice of the physician.



Education regarding safe storage and potential dangers of access to guns.

Changes in the manufacturing process to require guns to have safety mechanisms and to limit the "kill" capacity of guns by restricting the production of large capacity ammunition and assault weapons

Limiting access to guns through restricting the carrying of guns in public venues. Creating polices to limit the sale of guns to individuals who may be convicted felons or who have severe mental illness. Child access laws requiring parents to store guns safely.

Fig. 4. Public Health Triad Model

Only 6% claimed that they would ignore or be offended by such advice (Haught , Grossman & Connell, 1995). In a study by Albright & Burge (20 04) patients were randomly assigned to a control group where no gun counseling took place, verbal counseling only, or verbal counseling plus a brochure was disseminated. All education given focused on safe gun storage practices. Family physicians' brief counseling efforts made a significant positive impact in the gun storage habits of their patients. With a verbal or written recommendation, a significant improvement was observed in firearm storage. While there is evidence to suggest that counseling by healthcare and mental health providers may reduce gun access and promote safe storage many practitioners may not see this as there role and may not have received any professional training on how to discuss this issue with patients (Kinnison, Price & Telljohann et al., 2010; Price, Thompson & Khubchandani, nd). More training is needed for providers and additional information should be required during practicums or residency training programs (Coyne-Beasley & Theodore, 2006).

6. Conclusion

Guns have a significant impact on public health in the United States. Given the current toll of gun violence, it is clear that action is needed. While it has yet to be completely determined the extent to which all gun policies reduce gun violence and death, policies have the potential to reduce the number of firearm fatalities and injuries. While personal infringements are always a concern when enacting firearm policies there are occasions in which the "collective good" must outweigh individual perceived rights. A substantial number of the thousands of firearm homicides, suicides, and unintentional injuries should be able to be prevented without infringing too heavily on individual rights (Cook, Moore, Braga, 2001). It is only through alternative approaches such as education, policy implementation and evaluation that we can hope to develop confident conclusions about which strategies may reduce firearm violence and death. The goal of gun control programs over the next several decades should be to develop, implement, and evaluate specific gun control measures that can reduce firearm homicide, suicide, injury, and crime while preserving much of the legitimate uses of guns as possible (Cook et al., 2001). Both gun owners and nonowners express support for some gun control policies yet legislators at the state and the federal level continuously fail to enact such legislation.

7. Acknowledgement

The author would like to acknowledge Dr. James Price and Dr. Joseph Dake at the University of Toledo for their contribution to this work.

8. References

Albright, T. & Burge, S. (2003). Improving firearm storage: Impact of a brief intervention of firearm counseling by family physicians. *Journal of American Board Family Practice,* 16,40–6.

American Medical Association on Scientific Affairs. (1989). Firearm injuries and deaths: A critical health issue. *Public Health Reports,* March 111-120.

Annest, J., Mercy, J., Gibson, D., Ryan, G. (1995). National estimates of the nonfatal firearm-related injuries: beyond the tip of the iceberg. *Journal of the American Medical Association,* 273(22),1749-1754.

Ayres, I., Donohue, J. (2003). Shooting down the more guns, less crime hypothesis. In *Evaluating Gun Policy* Brookings Institute Press pp.287-341.

Black, D., Nagins, D. (1998). Do right-to-carry-laws deter violent crime? *Journal of Legal Studies*, 27, 209-219.

Blendon, R., Young, J., Hemenway, D. (1996). The American public and the gun control debate.*Journal of the American Medical Association,* 275 (22), 1719-1722.

Bonnie, R., Fulco, C., Liverman, C. (1999). Reducing the burden of injury: advancing prevention & treatment. Committee on Injury Prevention and Control, Institute of Medicine Report. National Academy Press. Washington, DC.

Boyd, J., Moscicki, E. (1986). Guns and youth suicide. *American Journal of Public Health,* 76, 1240-1242.

Brady Center. State firearm legislation. Accessed on September 23, 2011 at http://www.bradycenter.org.

Brent, D., Perper, J., Moritiz, G., Baugher, M. (1994).Suicide in affectively ill adolescents: a case-control study. *Journal of Affective Disorders,* 3, 193-202.

Bureau of Alcohol Tobacco and Guns. The Gun Control Act of 1968. Accessed on September 24, 2011 at http://www.atf.treas.gov/pub/fire-explo_pub/gca.htm.

Bureau of Alcohol, Tobacco, and Guns, U.S. Department of Treasury. Commerce in Guns in the U.S. February 2000. Washington DC.

Bureau of Justice Statistics. (2011). FBI Crime in the U.S. Report.

Bureau of Justice Statistics. (2001). National Crime Victimization Survey, 1993-1999.

California Department of Health Services. (1998). Residential Firearm Homicides in California, Report no. 3

Centers for Disease Control and Prevention. (2011) WISQUAR database. Accessed on October 17 , 2011.

Centers for Disease Control and Prevention. (2011) YRBSS database. Accessed on October 17 , 2011.

Centers for Disease Control and Prevention. (1999). Nonfatal and fatal firearm related injuries in the United States, 1993-1997. *MMWR*, l48 (45), 37-42.

Centers for Disease Control and Prevention/National Center for Health Statistics. (1997). Fatal Firearm Injuries in the United States 1962-1994. Accessed on October 17, 2011 at www.cdc.gov/ncipc/pub-res/firarmsu.htm.

Centers for Disease Control and Prevention. (1994) Deaths resulting from firearm and motor vehicle related injuries 1968-1994. *MMWR*, 43(3), 37-42.

Center to Prevent Handgun Violence. (1996). Denying handguns to prohibited purchasers: quantifying the Brady Law: executive summary. Accessed on March 23, 2003 at http://www.handguncontrol.org/gunlaw/bradylaew.htm.

Center for Responsible Politics. Gun Rights Recipients. Accessed on June 23, 2002 at http://www.opensecrets.org/industries/recipes.asp?Ind-Q13

Christoffel K. (1991). Toward reducing pediatric injuries from guns: charting a legislative and regulatory course. *Pediatrics*, 88(2), 294-305.

Coben, J., Steiner, C. (2003). Hospitalization for firearm-related injuries in the United States, 1997. *American Journal of Preventive Medicine,* 24 (1), 1-8.

Controller General of the United States, Report to Congress. Handgun control and effectiveness and costs. Washington DC. U.S. Government Printing Office, February 1978.

Cook, P., Braga, A. (2001). Comprehensive guns tracing strategies and investigative use of new data on guns markets. *Arizona Law Review,* 43 (2), 277-310.

Cook, P., Lawerence, B., Ludwig, J., Miller, T. (1999). The medical costs of gunshot injuries in the United States. *Journal of the American Medical Association*, 277 (28), 1853-1859.

Cook, P., Ludwig, J. (1997). Guns in America: National survey on private ownership and use of guns. National Institute of Justice Report.

Cook, P., Ludwig J. (1996). Guns in America: results from a comprehensive survey of gun ownership and use. Washington: Police Foundation

Cook, P., Molliconi, S., Cole, T. (1995). Regulating gun markets. *Journal of Criminal Law and Criminology,* 86 (1), 55-92.

Cook, P., (1991). The technology of personal violence, in *Crime and Justice: Annual Review.* 14, University Press

Coyne-Beasley., T & Theodore, A. (2006). Future physicians: The need for additional training in firearm injury prevention training. *Southern Medical Journal 99 (3) 198-199.*

Denno, D., Grossman, D., Britt, J. (1996). Safe storage of handguns. What do police recommend? *Archives of Pediatric and Adolescent Medicine,* 150, 927-31.

Duggan, M. (2001). More guns, more crime. *Journal of Political Economy,* 109,1086-1114.

Federal Bureau of Investigation. (2000). Uniform Crime Reports for the United States:2000. Washington D.C.: U.S. Department of Justice.

Federal Bureau of Investigation. (1996) Uniform Crime Reports for the United States:1995. Washington D.C.: U.S. Department of Justice.

Fingerhut, L., and Kleinman, J. (1990). International and interstate comparisons of homicide among young males. *Journal of the American Medical Association*, 263,3290-3295.

Frattaroli, S., Webster, D., Teret, S. (2002). Unintentional gun injuries, firearm design, and prevention: what we know, what we need to know, and what can be done. *Journal of Urban Health*, 79(1), 233-239.

General Accounting Office. (1991). Trauma care: Lifesaving system threatened by unreimbursed costs and other factors. Washington, DC, General Accounting Office 1991 Report.

General Accounting Office. (1991). Accidental shootings: many deaths and injures caused by guns could be prevented. Washington, DC, General Accounting Office 1991 Report.

Hardy, M., Armstrong F., Martin ,B., Strawn, K. (1996). A firearm safety program for children: they just can't say no. *Developmental and Behavioral Pediatrics,* 17, 216-221.

Haught K, Grossman D, Connell F. (1995). Parents' attitudes toward firearm injury prevention counseling in urban pediatric clinics. *Pediatrics*, 96,649-653.

Hayes, R., LeBrun, E., Christoffel, K. (1999). Missing in action: health agencies lack critical data needed for firearm injury and prevention. Accessed on September 20, 2011 at http://www.childmmc.edu

Hoyer, D., Kochanek, K., Murphy, S. (1999). Final data for 1997. National Vital Statistics Reports 47(19). Accessed on July 27, 2003 at
www.cdc.gov/nchs/releases/00news/final death97.htm

Kinnison, A., Price, J., Telljohann, S., King, K. & Thompson, A. (2010). Clinical Psychologists Risk Management Practices and Perceptions. *Journal of Community Health, 10,60-67.*

Ikeda, R., Gorwitz, R., James, S., Powell, K, Mercy J. (1997). Fatal firearm injuries in the United States, 1962-1994. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention.

Johnson RM, Coyne-Beasley T, Runyan CW. Firearm ownership and storage practices, U.S. households, 1992-2002. *American Journal of Preventive Medicine*, 2004: 27(2):173-82.

Kellerman, A., Westphal, L., Fischer, Harvard B. (1995). Weapon involvement in home invasion crimes. *Journal of the American Medical Association,* 273, 1759-1762.

Kellerman, A., Rivara, F. (1993). Gun ownership as a risk factor for homicide in the home. *New England Journal of Medicine,* 329, 467-472.

Kellerman, A., Rivera, F., Somes, G., Reay, D., Francisco, J. (1992). Suicide in the home in relation to gun ownership. *New England Journal of Medicine,* 327, 467-471.

Kellermen A. (1991). The epidemiological basis for the prevention of firearm injuries. *Annual Review of Public Health,* 12, 17-40.

Kellerman, A., Reay D. (1986). Protection or peril? An analysis of firearm related deaths in the home. *New England Journal of Medicine,* 314,1557-1560.

Khubchandani, J, Wiblishauser, M., Price, J. &Thompson, A. (2011) .Graduate Psychiatric Nurse's Training on Firearm Injury Prevention. *Psychiatric Nursing 25 (4)245-252.*

Kleck, G. (1996). Gun buy back programs: nothing succeeded like failure. In *Under Fire.* Police Executive Research Forum, Washington. pp. 29-53.

Kleck G & Gertz, M. (1997). The illegitimacy of one sided speculation: Getting the defensive gun use estimate down. *Journal of Criminal Law and Criminology, 87,1446.*

Koop, C., Lundberg, G. (1992). Violence in America: a public health emergency. *Journal of the American Medical Association,* 267, 3075-3076.

Krug, E., Powell, K., Dahlberg, L. (1998). Firearm deaths in the United States and 35 other high-and upper-middle income countries. *International Journal of Epidemiology,* 27, 214-221.

Leenars, A., Lester, D. (1994). Effects of gun control on homicide in Canada, *Psychological Reports,* 75, 81-82.

Loftin, C., McDowall, D., Wiersma, B., Cottey, T. (1991). Effects of restrictive licensing of handguns on homicide and suicide in the District of Columbia. *New England Journal of Medicine,* 325, 1615-1620.

Ludwig,J., Cook, P. (2003). *Evaluating Gun Policy Effects on Crime and Violence* Brooking Institute Press, Washington DC.

Ludwig, J., Cook, P. (2000). Homicide and suicide rates associated with implementation of the Brady Handgun Violence Prevention Act. *Journal of the American Medical Association,* 284 (5), 585-591.

Ludwig, J. (1998). Concealed gun carrying laws and violent crime: evidence from state panel data. *International Review of Law and Economics,* 18 (3), 239-254.

Lott, J., Mustard, J. (1997). Crime, deterrence, and right-to carry concealed handguns. *Journal of Legal Studies,* 26, 1-68.

Max, W., and Rice, D. (1993). Shooting in the dark: estimating the cost of firearm injuries. *Health Affairs,* 12 (4), 171-185.

May, J., Hemenway, D., Hall, A. (2002). Do criminals go to the hospital when they are shot? *Injury Prevention,* 8(3), 236-8.

McDowall, D., Loftin, C, Wiersma. (1995). Easing concealed firearm laws: effects on homicide in three states. *Journal of Criminal Law and Criminology,* 86(1), 193-206

McGonial, M., Cole, J., Scwhab, C. (1993). Urban firearm deaths: A five-year perspective. *Journal of Trauma,* 35, 532-537.

Miller, M., Azrael, D., Hemenway, D. (2001). Firearm availability and unintentional firearm deaths. *Accident Analysis and Prevention,* 33, 477-484.

Nelson, D., Grant-Worley, J., Powell, J. (1996). Population estimates of household firearm storage practices and firearm carrying in Oregon. *Journal of the American Medical Association,* 275, 1744-1748.

Pierce, G., Bowers W. (1992). The Bartlett-Fox Gun Law's short-term impact on crime in Boston. *Annals of the American Academy of Political and Social Science,* 708,120-137.

Price, J., Thompson, A., Khubchandani, J &Wiblishauser, M. & Dowling, J. Perceived roles of emergency department physicians regarding anticipatory guidance on firearm safety . Unpublished manuscript.

Price, J., Thompson, A. & Dake, J. (2004). Factors Associated with State Variations in Homicide, Suicide, and Unintentional Firearm Deaths. *Journal of Community Health*,29 (4),271-283.

Reuter, P., Mouzos, J. (2003). Australia: a massive buyback of low-risk guns. In *Evaluating Gun Policy* Brookings Institute Press, Washington. pp.121-156.

Rosenfeld, R. (1996). Gun buy backs: crime control or community mobilization. In *Under Fire.* Police Executive Research Forum, Washington. pp1-28.

Seiden, R. (1977). Suicide prevention: a public health/health policy approach, *Omega,* 8, 276-275.

Shah, S., Hoffman, R., Wake, L., Marine, W. (2000). Adolescent suicide and household access to guns in Colorado: Results of a case-control study. *Journal of Adolescent Health,* 26, 157-162.

Smith, T., Smith R. (1995). Changes in firearm ownership among women: 1980-1994. *Journal of Criminal Law and Criminology,* 86, 133-149.

Smith, T. (2001). National Gun Policy Survey of The National Opinion Research Center: Research Findings. National Opinion Research Center. University of Chicago. December 2001

Task Force on Community Preventive Services (2005). Guide to community preventive services. Guns law. Accessed on October 19, 2011 at www.thecommunityguide.org/violence/firearmlaws.html

Vassar, M., Kizer, K. (1996). Hospitalizations for firearm-related injuries: A population-based study of 9562 patients. *Journal of the American Medical Association,* 275,1734-1739.

Vernick, J., Webster, D. Hepburn, L. (1999). Effects of Maryland's law banning Saturday night special handguns on crime guns, *Injury Prevention,* 5, 259-263.

Vernick, J., Meisel, Z., Teret, S. (1999). I didn't know the gun was loaded: An examination of two safety devices that can reduce the risk of unintentional firearm injuries, *Journal of Public Health Policy,* 2, 427-440.

Vernick, J., Teret, S., Webster, D. (1997). Regulating firearm advertisements that promise home protection, *Journal of the American Medical Association,* 277(6), 1391-1397.

Villaveces, A., Cummings, P., Espitia, V., Koepsell, T., McKnight, B., Kellerman, A. (2000). Effect of a ban on carrying guns on homicide rates in 2 Columbian cities, *Journal of the American Medical Association,* 283(9), 1205-1209.

Webster, D., Vernick, J., Hepburn L. (2002). Effects of Maryland's law banning Saturday night special handguns on homicides. *American Journal of Epidemiology,* 155 (5), 406-412.

Webster D, Vernick J, Hepburn L. (2001). Relationship between licensing, registration, and other gun sales and the source of crime guns, *Injury Prevention,* 7(3), 184-189.

Webster, D., Starnes, M. (2000). Reexamining the association between child access prevention gun laws and unintentional shooting deaths of children. *Pediatrics,* 106(6), 1466-1469.

Webster, D., Vernick, J., Ludwig J., and Lester, K. (1997). Flawed gun policy research could endanger public safety. *American Journal of Public Health,* 87, 918-921.

Webster, D., Gainer, P., Champion, H. (1993). Weapon carrying among inner-city junior high school students: defensive behavior vs. aggressive delinquency, *American Journal of Public Health,* 83, 1604-1608.

Weil, D., Knox, R. (1996). Effects of limiting handgun purchases on interstate transfer of guns. *Journal of the American Medical Association,* 275(22), 1759-1761.

Wintemute, G. (1996). The Relationship between Firearm Design and Firearm Violence: Handguns in the 1990's. *Journal of the American Medical Association*, 275, 1749-1753.

Zimring, F., Hawkins, G. (1987). *The Citizens Guide to Gun Control*. New York: Macmilian.

25

The Public Health Intervention of Skin Care for All: Community Dermatology

Terence J. Ryan[1], Steven J. Ersser[2] and Lucinda Claire Fuller[3]
[1]*Task Force of the International Society of Dermatology, Oxford University*
[2]*International Skin Care Nursing Group, University of Hull,*
Faculty of Health and Social Care,
[3]*International Foundation for Dermatology, Chelsea and Westminster Hospital, London*
UK

1. Introduction

So often said, “The skin is the largest organ of the body and the most visible.” Less often emphasised is *Skin Failure,* a public health problem of great magnitude that is global and a major precursor of long term disability and non-participation.

Skin failure manifests as; 1) the loss of barrier function between the body and a threatening environment; 2) the inability to manage overheating or excessive cooling; 3) sensory impairment so that there is itch, pain or numbness, and 4) the disfigurement and consequent faulty communication of the ‘look good feel good factor’.

Skin failure is managed by physician belonging to the profession of Dermatology that has had several years of additional training in the diagnosis and management of skin diseases. There are also wound healers, lymphoedema managers,employees of burns units, and custodians of neglected tropical skin diseases such as, for example, leprosy, leishmaniasis,onchocerciasis, lymphatic filariasis, yaws and Buruli ulcer, which must be members of an alliance to provide skin care and promote skin health for all.

The perception of Public Health is that the dermatologist is one who treats individuals with skin disease and who markets good looks served by cosmetic interventions. It fails to recognise that there is significant branch of Dermatology named Community Dermatology which promotes ‘*Skin care for all*’ (1). The mission ‘Healthy Skin for all’, an offspring of the WHO *Health for All* (2)*,* adopted first of all by The International League of Dermatological Societies and later by some of its members such as the British Association of Dermatology, is Utopian. *Skincare for all* is more achievable. It focuses on low cost interventions, relying on self help and low technologies such as washing, as well as knowledge of how to examine the skin and the acquiring of the skills of the diagnosis of physical signs in the skin.

Recent discussions at the WHO have initiated the view that *Health for All*, which is inclusive of wellbeing, would be more appropriately renamed if it were to focus on capacity to adapt and self-manage (3). The end point of treating skin failure would be a disfigured person who is coping well with self-management and feels healthy.

Fig. 1. A Photograph that has become a much used *logo* of the *International League of Dermatological Societies* to illustrate *Healthy Skin for all(2).* Those who treat must have access, (the bike) essential drugs and devices, (in the accompanying portable case) and knowledge(appropriate education).

The nature of community dermatology, and its clear capacity to benefit, has been highlighted by the International Society of Dermatology in 10 articles published in the May 2011 edition of the International Journal of Dermatology and in a CD of 44 articles prepared for the World Congress of Dermatology held in South Korea in the same month www.intsocderm.org. The capacity to benefit within the skin care field is dependent not only on the medical profession but on an effective response from the nursing service, the largest group of health professionals worldwide. This situation is directly relevant also to both the prevention and management of chronic wounds and Neglected Tropical Diseases worldwide, given the major involvement of nurses in routine skin care and maintenance of skin integrity.

2. The background to community dermatology

It is a significant advance that there has been recent effort by the WHO to record the burden of diseases and there is a sub-committee tasked with recording skin disease (discussed by Hay et al 2011 at the World Congress of Dermatology in Korea).Hay and Fuller (2011)(4)state that the questions that must be asked are

1. What and how common are the main local diseases?
2. What is their impact on those affected by them in terms of disability, reduction in quality of life, and cost?
3. What obstacles prevent people from seeking remedies?

These range from lack of understanding of the potential for treatment, tolerance of disease, distance to health center, to cost of treatment.

Some countries such as the UK(2009) (4) and USA(2006) (5), have much improved the accuracy and approach to collection of data on skin disease and its assessment and expression as a burden with needs to be met. Assessment and management in resource poor environments are now improving but still inadequately funded. In the developing and emerging countries, the education needed to ensure skilled examination and recognition of skin conditions is almost everywhere lacking. But with imported expertise some thorough analyses of the prevalence of skin disease in rural areas have been done, viz: Ethiopia,(6,7)Nepal,(8) Mali)(9) Tanzania,(10) Indonesia(11) and Brazil(12.) These show a high prevalence of skin disease, usually exceeding 50% and sometimes when there are endemic conditions such as scabies, onchoceriasis, tinea capitis or lice, as much as 80% of the population, For instance, Walker's study in Nepal(8) used data from a survey conducted in five villages and found an overall point prevalence of skin diseaseof 62.4%. Superficial fungal infections accounted for 20%, but acne, melasma and eczema were among the top five diagnoses.

Fig. 2. A rural clinic in Ethiopia for women with podoconiosis, a condition of lymphoedema due to not wearing shoes where the soil is irritant(13). In this clinic for foot inspection,better care of the feet with significant improvement in the skin is awarded with microfinance for job development.

Another medical problem endemic in Ethiopia (Figure2) is podoconiosis (13), a chronic form of lymphedema associated with exposure,through not wearing shoes, to certain soil types, where in one village 5.8% of those over 50 years old had this problem.

In Gibbs' study of two villages in Tanzania,(10) there was a lower prevalence rate for skin disease (28.6%), but again over 70% of these patients had skin infections.

In a survey in Assuit (14) Egypt a higher level of skin disease, with 87% of those examined having at least one skin problem being recorded. This study included individuals with pediculosis capitis, which had a prevalence of 19.37%. In addition, 10% had bacterial skin infections and 16% fungal infections. Eczema prevalence was also high at 19%.(14) By contrast, the survey carried out in Sumatra, Indonesia(11) showed a lower prevalence of skin problems. In children and adolescents, 28.2% of the population studied had skin disease although, as is usually found, fungal and bacterial infections were the most common diseases seen. It is children under the age of 15 years that account for most of the diseases affecting the skin in developing countries. For instance, Mahé's survey in Mali (9) showed that 34% of children had one or more skin problems, with infections predominating. Any such study will be aware that the world's population is dominated by children. A study of skin disease in 12,586 Indian school children ranging in age from 6 to14 years showed an overall point prevalence of skin conditions of 38.8%. Of these, 30% had only one skin disease, 6% had two and 2.7% had three skin problems.(15) The most common of these were infections (11.4%), but pityriasis alba (8.4%), eczemas (5.2%) and infestations (5.0%) were also common. Bechelli in Brazil (12) found a prevalence of skin disease in the villages surveyed that ranged from 20 to 87% in children, aged 5–16 years,living in different municipalities in Acre state, Brazil. Again, infestations and infections were the dominant conditions; for instance, the prevalence of pediculosis capitis exceeded 50% in most areas.

A common theme of all these studies is that the prevalence of disease is often affected by the prevalence of parasitic infections, such as scabies. A key point is that parasites carry other infections and scratching exposes the skin to additional infections and attracts flies. Some medicaments such as honey attract other insects such as ants..A special concern is the prevalence of streptococcal infection in those populations where scabies is prevalent. For instance, in Northern Australia amongst aboriginal children, scabies often affects over 50% of the population, and pyoderma and infected sores with staphylococcus and streptococcus affect 70%.(16). Quality of life can be seriously affected, clinics can be overwhelmed, rarer but very important diseases given less attention, and family money spent on ineffective remedies.

It is not only dermatology medical professionals that lead advocacy for a public health approach to these skin diseases,but, as well, there are other health professionals taking an interest. The International Skin Care Nursing Group (ISNG), initiated in the UK, affiliated both to the International Council of Nursing and to the International League of Dermatological Societies (ILDS), is giving additional nursing leadership (17). From its earliest days members of its board have acted as a catalyst by visiting departments in India, Mexico,Nepal and Africa. In Guatemala the International Foundation of Dermatology with some help from the ISNG focused on teaching 255 American Indian nurses working in rural areas. This was 10 days of instruction into skin care over a 5-year period (18). In the last few years in countries such as India and Nepal, where the nurse is predominantly female and dermatology nursing is not well developed, the ISNG has begun to exert an influence by demonstrating how important it is to have a profession providing skin care that uses hands-on low technology in skin care delivery. The nurse becomes an educator and counsellor, with more time than the rare dermatologist, available to teach the patient self-help. In the Community Dermatology projects led by Dr Roberto Estrada Castanon in rural Mexico, the nurse acts as a "sentinel", diagnosing and managing patients in their rural homes. In Manaus, Brazil, in the Dermatology Clinics and the Institute of Tropical Medicine, all attendees are first seen by a dermatology- trained Triage nurse, who selects whether they should be returned to primary care, tested for malaria, admitted to a ward or referred for a skin consultation.

The significant lack of dermatologists worldwide, most of whom are based in the secondary hospital sector, means that expertise in skin care cannot always be delivered to those who need it. A paper written by the International Skincare Nursing Group (18) outlines initiatives by nurses to do this and examines the development of nursing within the dermatology field- and the need to cascade dermatological expertise from specialist to generalist community based health workers. It logically follows that there is a need to adopt a strategic approach that identifies the educational needs of nurses, harnesses the appropriate expertise, shares good practice and operates in close inter-professional collaboration with dermatologists. Work led by the *International Skin Care Nursing Group* (ISNG) has sought to stimulate and develop the capacity of nursing to respond to these widespread needs through promoting service delivery models that operate interdependently with dermatologist-led care. Such health professionals are well placed to work with community workers in resource poor countries and manage core clinical issues, including skin barrier protection and maintenance and the management of chronic wounds, applying preventative and educative strategies where possible. Evidence of such work is seen in South Africa with the development of a new Postgraduate Diploma in Dermatological Nursing to adequately prepare a cadre of nurses who can work in peripheral community based clinics addressing issues of skin and wound integrity. To realise this potential there is a need for the development of inter-professional team working, support for nursing development and education and identification of service delivery models that effectively integrate and complement their expertise. In this context diagnosis is based on focusing on the common diseases. The most frequent of these are usually infections such as impetigo, cutaneous fungal infections or parasites such as scabies. Attention to better diagnosis and the availability of essential drugs must include awareness and reminders of diseases such as the Neglected Tropical Skin Diseases. Reducing the prevalence of ineffective remedies often bought at the road side, is also a part of the programme.

2.1 Gender and the provision of Skin care

Those involved in skin care, including dermatologists, allied health professionals, (principally nurses) or most often a family member, are more likely to be female than male.

Women are the largest users of health services, including alternative and complimentary medicine.

The empowerment of women to improve their effectiveness has the backing of an organisation founded in the memory of Maria Duran, a Columbian Dermatologist who helped to organise in Turkey the first meeting on the Advancement of Women in

Dermatology (19). Dermatology is a profession in which women are successful and numerous, and much past criticism of the undervaluing of women in professions and in universities has been influenced for the better by the experience of dermatology. Thus, especially in the integration of an immigrant population, dermatology has an important role to play. It continues to be a model of how women in social class 1 cope with a professional career. The late Dr Turkan Saylan, Professor of Dermatology and Leprology at Istanbul University, stated that the solution to ''Advancement of Women'' is humaneness rather than laws, lore or religion(20). It was the International League of Dermatology Societies (ILDS) and International Society of Dermatology that initiated the concept that dermatology could provide guidelines for equal partnerships between men and women to counteract the male dominance of many codes of practice taught by religions. The laws that govern health and hygiene are mostly written by men.

Dr Rokea A el-Azhary, of the Mayo Clinic, in a presentation at the Cairo International Congress of Dermatology in 1996, stated that the change in the ratio of women to men in dermatology would bring with it the necessary gender choice for the patient, a perception of greater kindness, better communication as well as greater empathy.

The setting up of vulva clinics in Dermatology departments (21)has led to a model of care that is neither based on sexually transmitted infections nor on the surgical approach of gynaecology to genital skin disease, such as in the management of lichen sclerosus et atrophicus. Many publications from women dermatologists now show that previously incurable conditions can now be better diagnosed and treated. In many countries, social norms dictate that female genital skin conditions be managed by female dermatologists and, indeed, many female patients prefer this. Measurement of quality of life has shown how miserable the lives of women suffering from genital skin disease have been (Finlay in CD intsocderm). Dermatology, through its knowledge of the handicap of disfigurement, can do a lot to seek ways to remedy this, and to raise self-esteem and reduce perceptions of having an appearance which is unwelcome.

Cultural concepts of beauty that are almost entirely to be found in the skin, affect marriage. Young girls must be given more confidence to delay marriage and a greater choice in their life partners. There has been improvement in maternal and child mortality through better diagnosis of some skin conditions in pregnancy. In future an aim is to ensure girls are as well nourished and well dressed as boys and, in a threatening climate, that they are as protected as boys. The Albino project of the Regional Dermatology Training Centre (RDTC) in Tanzania is a model for gender equality both for the treatment of skin cancers and for ensuring the educational needs of girls as much as boys.(22) Currently of major concern in Tanzania there is a market for the selling of body parts of those affected by albinism for witch craft and girls are most vulnerable.

Circumcision of males is currently advocated for protection against HIV/AIDs but better education on genital hygiene has a place in management of the foreskin. Female genital mutilation is a major concern and is discouraged by a global drive to reduce this practice often led from Dermatology.Thus Morrone a dermatologist on behalf of the International Society of Dermatology, wrote ''Stop female genital mutilation: an appeal to the international dermatologic community''.(23) At the RDTC in Tanzania, much support has been given to the drive against this procedure and to the education of health care professionals on how to discourage this practice.

2.2 The regional dermatology training centre: A flagship for community dermatology

In resource poor countries, in recent years, greater effort by leading organisations governing dermatology (The International Society of Dermatology, the ILDS and The International Foundation for Dermatology, IFD) resulted in the creation of the Regional Dermatology Training Centre in Tanzania, RDTC (www.ifd.org); this has produced 200 dermatology graduates serving 12 English speaking countries of Sub-Saharan Africa which were without any hospital or university advice on the skin except for the occasional visitor from the developed world. It initially focused on the development of Allied Health Professionals but the number of Specialist Doctors, MDs, in training is increasing. The two year curriculum includes skin diseases, STI and Leprosy. The students also receive instruction on wound management and lymphoedema. Another programme based in Mali, West African and French speaking, was a one day training of all health centre personnel to recognise and manage the three commonest conditions, resulting from bacterial, fungal and parasitic invasion of the skin: namely impetigo, superficial fungus and scabies. This training prevents the health services being overwhelmed by common diseases and allows the health worker to give more attention to more serious skin conditions such as leprosy, or wounds and burns. Other programmes now exist in Mexico, Mali, Patagonia and Cambodia (see the CD www.intsocderm.org) all administered or supported by The IFD from Willan House,4 Fitzroy Square London, the HQ of the British Association of Dermatology.

In the field of wound healing the curriculum includes improving techniques of making wounds; incision and suturing for example after biopsy or excision of early skin cancers and avoiding disfiguring scars. In resource poor regions effective low technology interventions are promoted including the use of herbal remedies such as honey and biosurgery employing maggots to aid wound healing.

Information Technology, including telehealth, is advancing in expertise and becoming more available as mobile phone technology and distribution expands. (See Schmidt Grendlemier in CD, www.intsocderm.org) It is used by all levels of health care, inclusive of traditional health practitioners of rural Africa. The skin is easily photographed and easily biopsied. Samples can be posted and histopathological analysis sent back to guide the management of skin lesions (see Bertramelli in CD, www.intsocderm.org). It is increasingly common practice to teach Allied Health Professionals to biopsy the skin using disposable punches. One author emphasised that honey was an excellent transport material even for histopathology needing immunoflourescence (24)

The morbidity and health issues facing mobility of populations forced to move due to strife, climate change or economic migration has been extensively studied and supported by the dermatologist Aldo Morrone (CD www.Intsocderm.org) through the establishment of the National Institute for Health, Migration and Poverty, in Rome.There are 214 million international migrants and 740 million internally displaced people. The immigrant, even if without skin ailments before migrating, is very likely to develop such during a difficult migration. In a paper in press, the human right to have access to skin care and the legal duty of the carer to provide it is emphasised(25)

Dermatology also overlaps and embraces venereology-sexually transmitted infections including HIV/AIDS. These conditions both present frequently in the skin and resulting therapy can commonly produce adverse skin reactions. Tissue viability may also be generally impaired and for example skin grafting more often fails in immunocompromised patients.

Skin Care for ALL is difficult to achieve when considering the challenges of climate change. (26). Those who care for the skin have developed expertise in the management of hazards from sun, flood and cold exposure. Programmes to protect the skin from the hazards of climate are well developed in Australia with their *Slip slap slop* sunscreen campaign, in Tanzania there is the Albino programme and along the western mountain ranges of the Americas there is protection of the American Indian genetically prone to Polymorphic Light Eruptions. Many conditions such as these which are influenced by climate may get worse unless the experience of skin carers is fully appreciated and funded.

Skin lightening with often harmful depigmenting agents containing strong steroids or mercury, has become a significant public health problem in Asia and Africa adding to the disorders of pigmentation such as albinism and vitiligo an increasing burden in primary health. It is a consequence of the perception that to be lighter implies a higher social class less burdened with outside work and better off financially.

2.3 Social marketing of self-esteem and 'look good feel good factor'

Much of the advice to patients and the support given to self-help groups, divorces beauty from market forces, advising on lifestyle and removing from their vocabulary being contagious or unclean.

Dermatology textbooks addressing social and sexual communication help to explain how display is used not only for sexual purposes, but in conflict behaviour to establish territorial rights or social dominance.(27) Social marketing (28) of concepts such as "Black is Beautiful" or "Natural is Beautiful" is important in reducing the use of skin lightening creams containing harmful agents such as strong steroids or mercury. Those who provide or receive these rely entirely on dermatology to make them safe, particularly from adverse skin reactions. Many cultures use cosmetics or tattoos to enhance communication. Often too there is a request for their removal with the passing of time.

The importance of hair as an adornment, especially in women, too much or too little, also requires advice from dermatologists.

There is a great increase in the demand of teenagers for tattoos,and body piercing followed by a more adult demand for their removal at a high cost. These are initiated by a desire to raise self esteem. The take up of smoking in adolescence is also part of the same need and even graffiti can be shown to be a 'marking' of the environment as part of impressing ones identity. The Community Dermatology Task Force believes that understanding these issues, and anything that encourages the carer of the skin to understand the need and to develop interventions to raise self esteem, are key to managing wellbeing. People with disfigurements due to skin ailments, or having selected a pattern of adornment no longer fashionable or having scarring and consequences of burns, in particular, need help and support by councillors well trained in managing communication skills. They can be helped by organisations who specialise in bringing out skills as an alternative to good looks.. When severely disfigured one may have to learn to walk into a room of people having relearned how to introduce oneself by both the loudness of one's voice and the strength of one's hand shake.

All such aforementioned programmes are a development of public health interventions, but they will succeed only if distributed alongside other public health programmes. Indeed Western biomedical practice cannot intervene successfully without awareness and a degree of partnership and effective integration with other systems of medicine. These include Indian and Chinese systems of medicine as well as Traditional Health Practice in Africa (29). There are many more Traditional Health Practitioners(THP) than biomedically trained practitioners. They are a huge work force with potential for delivering skin care sustainably and locally available. In managing a problem such as for example snake bite there are interventions they should not do, such as scarification and the use of the tourniquet. They can identify the snake. They can manage shock due to fear. Increasingly some urban based THPs even use a mobile phone and in future communication with experts from THPs in rural areas may be the custom, even showing over great distances a photo of the snake to antivenom centres and seeking advice whether venom is needed and available. The THP can be helpful in disease control, as for example managing HIV/AIDS or adding to the work force for elimination of disease campaigns as well as morbidity control. Well established written traditions, such as Indian or Chinese medicine are important collaborations if *Skin care for all* is to achieved. For example the thousands of years old traditions of Ayurvedic Medicine and Yoga have,inspite of initial biomedical opposition, been shown to be effective low cost management of lymphoedema due to filariasis in the villages of India.(30)

All programmes need to be patient-centred, to ensure effective patient engagement to uphold treatment adherence and empower people to self-manage. Even in community management, awareness of individual needs has to be preserved, but recognised as a part of the community profile. Where possible the self-management dimension for care should be supported through planned education and encouragement to enhance treatment adherence and prevention of skin barrier deterioration; nurses have a key role here. The extent to which the absence of skin care leads to morbidity and indeed at times mortality (31), should be emphasised through educational initiatives. Morbidity needs the back-up of formal *quality of life* scoring so that the burden is fully realised and the basis of community participation and income generation can be worked upon (32).

3. The future of community dermatology

The International Society of Dermatology Task force for '*Skin Care for ALL: Community Dermatology*', has set out the scope and range of the public health programme needed. Following on from this background work descriptions of interventions, both by addressing the service need in terms of workforce and materials required, must be acted upon.. The Task force is preparing an ABC of *how to do it* procedures to promote community dermatology. The agenda and curriculum has to be shared with many partners in an alliance to achieve *Skin Care for All*. Such an alliance will work with those organisations managing wounds, burns and lymphoedema and will work with the World Health Organisation to contribute to its interventions on skin care for the Neglected Tropical Diseases Programme. Many of these have a work force to deliver drugs to whole populations to break transmission from vectors and between humans. They all need backing with programmes of morbidity control. Leprosy is long a prototype for the Stigma of skin disease. Hopes for its elimination as a public health problem have not been fulfilled. There are still far too many new cases.Often these are multibacillary, have advanced disability and are seen in children. Such worrying aspects of this disease,particularly in India and Brazil, are made more serious by the widespread disbanding of Leprosy's vertical programme and therefore of its experienced workforce in the belief that the battle had been won. Programmes from Dermatology such as that at the Regional Dermatology Training Centre in Tanzania, which is a WHO collaborating centre for leprosy as well as Dermatology and STI, are insufficiently utilized within a framework of leprosy led control programmes.

Podoconiosis has recently been added to the list of Neglected tropical Diseases. This is a condition illustrated above (Fig2) for which active skin care and the introduction of footwear has the potential to benefit more than one million Ethiopians. Current simple NGO projects are already helping thousands of patients (see Davey 2011 CD www.Intsocderm.org). When contributing to a lymphatic filariasis programme in Kerala S India, the authors of this chapter, acting as mentors, have witnessed a significant advance in morbidity control through the integration of conventional medicine and traditional effective practices such as massage and yoga. With many millions of persons in India having grossly enlarged limbs, the pilot study of some 2000 limbs has revealed an effective response to integrated medicine, making use of the availability in villages of self help interventions using yoga and herbals (30).

Onchocerciasis elimination has focused on the use of the drug Ivermectin to eradicate the filarial infection but the needs of a heavily scratched skin covered in bites from black flies require a concomitant planned programme of skin care. Leishmaniasis is another common disease resulting in disfigurement which is greatly in need of techniques to prevent or treat scarring.

Buruli ulcer has recently benefited from improved instruction on the technology of wound healing. It was the WHO's need for wound healing expertise for Buruli ulcer that lead to the formation of the World Alliance for Wounds and Lymphoedema Care (WAWLC) and the publication of a White Paper (a document prepared by experts for advice to Governance)in the form of a Handbook is now on the shelves at WHO. (34) It simplifies management of chronic wounds down to five interventions and lymphoedema to three. Table 1

Wound healing

1. Manage systemic conditions such as diabetes, anaemia, HIV AIDS
2. Protect the wound from trauma such as sustained pressure
3. Promote a clean wound base and control infection
4. Maintain a moist environment
5. Control periwound oedema

Lymphoedema

1. Enhance movement of lymph
2. Reduce overload from the venous system
3. Reduce overload due to inflammation from entry points for irritants and bacteria.

Table 1. The minimal requirements for Wound healing and Lymphoedema

Fig. 3. Pyramid distribution of advice provided by carers of the skin must be both to Governance above as well as to practitioners below.

Experts working within the framework of a pyramid of care Fig 3 with the most needy at its base and governance at the top must give equal attention to both. At the bottom there is a need to provide locally available, sustainable, low cost care, mainly through primary care initiatives. The support of the nursing profession is crucial to this service delivery model. In the pyramid of care the Task force for skin care sits in the middle levels often collaborating with university and tertiary hospital care. For the management of prevention or first aid tertiary care extends downwards to the health centre, Traditional Health Practitioner, and to self help. For this to happen it is perhaps even more important that time is put aside for advocacy to extend upwards to those who govern health provision. So far this has resulted in little understanding of the significant role the carer of the skin plays in health care in general.

3.1 Wounds burns and Lymphoedema

Throughout the development of the Taskforce for *Skin Care for All: Community Dermatology* there has been emphasis that this is a programme extending beyond dermatology to include Wounds, Burns and Lymphoedema. This leads us into the public health territory of *Global Injury Epidemiology*. Injury is a leading cause of death, highest in middle income and lowest in lower income groups. While acute injury such as the head injury of a motor bike accident resulting in death cannot be included as part of the remit of the Task Force for skin care for all, trauma resulting in chronic non healing skin wounds should be. As mentioned above in reference to a paper by Escadon et al(31)these, such as pressure ulcers in the paraplegic, can lead to death and add to the little emphasised statistic of mortality due to skin conditions. Its prevalence is underestimated in data describing the scale of the injury problem.

The discussions that can be read in the literature of classification and coding of injury tend to focus on anatomical site such as head, trunk or limb but an *Injury Severity Score* does separately identify *external*, meaning the skin. The literature relating to injury includes much that is written about the economic and social burden. The care of the person living with an amputation stump, the pressure ulcer in the paraplegic and of course burns are just a few examples. Preventing and managing the consequences of injury becoming a top cause of loss of disability adjusted years falls commonly to the carer of the skin.

The burned patient is an appropriate subject for a Task force providing care for people with skin conditions that arrive in the primary care health centre or requiring self help. This hardly needs further emphasis, yet a glance at the headings so far illustrating how education provided by carers of the skin has the capacity to benefit, will reveal the following: - the greater susceptibility to burns of the skin without sensation, the importance of clean water availability, the need to avoid inflammable clothing and furnishings, child to child programmes to prevent the youngest approaching the fire place, and the whole question of disfigurement restoration and the look good feel good factor that permeates the management of all skin disease. These require environmental, engineering, legislative, and educational approaches requiring rules on inflammable clothing and bedding, as well as whole country teaching of first aid

If the definition of *Health* is to change, current debate may revise it to meaning the *capacity to adapt as an agency to effectively self-manage*, in which case those who care for the skin are well in the lead to contribute to *Health for all through* by their promotion of skin health.

4. Conclusion

Community Dermatology must pervade all health systems- making the correct diagnosis and intervening with available and affordable interventions, including health promotion and the prevention of the deterioration of the skin barrier and the development of chronic wounds. Data must be collected on prevalence of disease and access to interventions and the size of the work force, and Essential Drug Lists must be prepared. As an example of the globalisation of key interventions there is the Memorandum of Understanding between Procter and Gamble, the International Foundation for Dermatology and The International Skin Care Nursing Group to provide *Water Fit for Drinking* for washing by issueing very low cost water cleaning systems and adding glycerine as an emollient. This can be seen as a collaborative initiative promoting the lowest level of technologies used in skin care (35).

Public Health strategies have not been sufficiently inclusive of skin care. Public health inclusivity of common problems requiring skin care is largely unfunded while there is the receipt of billions of dollars for management of HIV/AIDS, malaria and TB. Donors have yet to be persuaded of the benefits of skin care possibly because of the lack of focus by public health professionals and perhaps a lack of appreciation of the public health relevance of skin health.

The task force for *Skin Care for All:* Community Dermatology can demonstrate in its recent publications its ‘capacity to benefit’. It will in the near future with greater clarity than ever before, explain how to approach the problem. But this will be a waste of potential achievement for solving very common public health needs unless a strategic approach is taken and the resources are provided to enlarge a work force that is skilled and equipped to aid effective self-management. An alliance of all who interested in skin care will be needed.

5. References

[1] Ryan TJ. The International Society of Dermatology's Task Force for *Skin Care for All* : Community Dermatology *International Journal of Dermatology* 2011;50: 548-551

[2] Ryan TJ. Healthy Skin for All.International Committee of Dermatology *International Journal of Dermatology* 1994;33:829-835

[3] Huber M, Knottnerus JA, Green L, Vander Horst H et al. (2011) Health :how should we define it.? *British Medical Journal* 343:235-237

[4] Schofield JF, Grindlay D and Williams H C. Skin conditions in the UK: a Health Care Needs Assessment. Centre of Evidence Based Dermatology, University of Nottingham, UK 2009

[5] Bickers DR, Lim HW, Margolis D, Weinstock, Goodman C, Faulkner E, Gould C, Gemmen E, and Dall T, The burden of skin diseases: 2004. A joint project of the American Academy of Dermatology Association and the Society for Investigative Dermatology J.Am.Acad.Dermatol 2006;55:490-500.

[6] Figueroa JI. Fuller LC. Abraha A and Hay RJ 1998. Dermatology in southwestern Ethiopia: rationale for a community approach. Internat. J. Dermatol. 37:752-758

[7] Leekassa R. Bizuneh E. Alem A. Fekadu A. Shibre T. Community diagnosis of common skin diseases in the Zay community of the Zeway Islands, Ethiopia. Ethiopian Med. J. 2005 43:189-195

[8] Walker SL. Shah M. Hubbard VG. Pradhan HM. Ghimire M. Skin disease is common in rural Nepal: results of a point prevalence study. Br J.Dermatol. 2008 158:334-338

[9] Mahé A, Thiam N'Diaye H and Bobin P.1997. The proportion of medical consultations motivated by skin diseases in the health centers of Bamako (Republic of Mali). Int J Dermatol 36:185-186

[10] Gibbs S 1996. Skin disease and socioeconomic conditions in rural Africa: Tanzania. Internat.J.Dermatol. 35:633-639

[11] Saw SM. Koh D. Adjani MR. Wong ML. Hong CY. Lee J. Chia SE. Munoz CP and Ong CN. 2001. A population-based prevalence survey of skin diseases in adolescents and adults in rural Sumatra, Indonesia, 1999. Trans Roy Soc. Trop.Med.Hyg. 95:384-388

[12] Bechelli LM, Haddad N, Pimenta WP, Pagnano PMG, Melchior, Jr E, Fregnan RC, Zanin LC, Arenas A: Epidemiological Survey of Skin Diseases in Schoolchildren Living in the Purus Valley (Acre State, Amazonia, Brazil). *Dermatologica 1981;163:78-93*

[13] Davey G Podoconiosis, non-filarial elephantiasis, and lymphology. Lymphology [2010, 43:168-77]

[14] Abdel-Hafez K. Abdel-Aty MA. Hofny ER. Prevalence of skin diseases in rural areas of Assiut Governorate, Upper Egypt. Int J.Dermatol. 2003, 42:887-892,

[15] Dogra S. Kumar B Epidemiology of skin diseases in school children: a study from northern India. Pediatric Dermatol. 2003 20:470-3

[16] Currie BJ. Carapetis JR. Skin infections and infestations in Aboriginal communities in northern Australia 2000. Austral.J.Dermatol. 41:139-43

[17] Ersser, S. J, Kaur V, Kelly, P, Langoen, A, Maguire, S, Nicol, N.H, Page B, and Ward, C. The contribution of the nursing service worldwide and its capacity to benefit within the dermatology field. *International Journal of Dermatology*, 2011;50: 582-589.

[18] Ryan T J (1994) Nurses complete their training in Chimaltengo. Int J Dermatol 33:526-7

[19] Murrell D, Ryan TJ, Bergfeld WF. Advancement of women in Dermatology *International Journal of Dermatology 2011;* 50: 593-600

[20] Saylan T. Prevention of causes of traffic in women and forced prostitution; approaches and actions. Report of the European Commission for Equality between Women and Men.EG/Prost 1990;91:3

[21] Ridley CM International Society for the Study of Vulvar Disease. J Reprod Med. 1993 Jan;38(1):1-4.

[22] McBride SR, Leppard BJ. Attitudes and beliefs of an albino population toward sun avoidance: advice and services provided by an outreach albino clinic in Tanzania.Arch Dermatol. 2002 May;138:629-32.

[23] Morrone A, Herzogova J, Lotti T. (1992) Stop female genital mutilation: and appeal to the International Dematologic community. In J Derm 41:253-263

[24] Ziena B,Honey Transport Abstract BAD annual meeting supplement

[25] Ryan TJ, Morrone A,Hay R, Fuller LC,Naafs B, Sethi A,Guidelines on the role of Skin care in the Management of Mobile populations. Based on a workshop Rome 24th-25th March 2011 International J Derm in Press

[26] Anderson LK.Global climate change and its dermatological diseases. *International Journal of Dermatology 2011;*50: 601-607

[27] Ebling FJG.Socio-sexual communication.In Rook A, Wilkinson DS,Ebling FJG eds Textbook of Dermatology.Oxford: Blackwell Scientific1991 p188

[28] Karelas GD. Social Marketing self-esteem: a socio-medical approach to high-risk and skin tone alteration activities. *International Journal of Dermatology 2011;* 50:590-592

[29] RyanTJ, HirtH-M, Willcox M.Collaboration with Traditional Health Practitioners in the provision of skin Care for all in Africa. *International Journal of Dermatology* 2011;50:564-570

[30] Narahari SR, Ryan TJ, Bose KS et al. Integrating modern dermatology and Ayurveda in the treatment of vitiligo and Lymphoedema in India. *International Journal of Dermatology2011;*48: 310 334

[31] Escandon J, Vivas AC, TangJ, Rowland KJ, Kirsner RS, High Mortality in Patients with Chronic wounds. *Wound Repair and Regeneration*. 2011;9:526-528

[32] Basra MKA, Fenech R, Gatt RM, Salek MS, Finlay AY. (2008) The Dermatology Life Quality Index 1994-2007: a comprehensive review of validation data and clinical results. *British Journal of Dermatology 159: 997-1035.*

[33] Macdonald J and MJ Geyer. Wound and Lymphoedema Management WHO/ HTM.NTM/ GBUI.2010

[34] Brooks J. Water fit for drinking is fit for washing wounds! A case study at a Ugandan Hospital. *Community Dermatology 2011;* 7: 17-32

26

Addressing Asthma from a Public Health Perspective

Adam Davis and Mindy Benson
Children's Hospital and Research Center Oakland
USA

1. Introduction

Asthma's prevalence, treatment costs, an impact on quality of life make it one of the most pressing global public health issues, especially in developed nations (Asher et al., 2006). Despite tremendous progress over the past 3 decades in the quality, variety, and efficacy of asthma treatment and prevention options, improved coverage by medical managed care organizations, and greatly improved air quality, the prevalence, morbidity, and related costs of asthma have not decreased over this same time period. (Moorman et al., 2007).

Many well-designed clinical and prevention-oriented interventions in various settings have been able to demonstrate short-term reductions in asthma morbidity for individuals and small groups of participants in controlled settings (Coffman et al, 2008; National Heart Lung and Blood Institute 2007). However, these interventions often fail to have an impact on a larger scale. There is still a compelling need to identify how best to reduce the burden of asthma for a defined population, such as a city, state, neighborhood, or membership of a health plan.

The population perspective is the hallmark of public health (Glasgow et al., 1999). Interventions that impact populations are important because they have the potential for reducing societal costs of treating the disease and promoting the health of a larger number of individuals (McNairm et al., 2007). The questions and issues that go into planning interventions on populations is not simply the "scaling up" of clinical or small-scale interventions that have demonstrated efficacy in controlled settings, but is fundamentally different.

The preliminary steps that commonly constitute public health program planning--conducting a needs assessment, constructing a logic model, and building a coalition of stakeholders--are well documented (Centers for Disease Control and Prevention 2010a, 2010b; Institute of Medicine 1988). However, guidance for directing limited resources to affordably maximize the effect on population-level measures of asthma is extremely limited. To achieve sustainable population-level impacts interventions must not only be effective, but also have high reach, wide adoption, and affect change among patients and institutions over the long term (Glasgow et al, 1999). Public health's challenge is to have an impact on large numbers of individuals with asthma using effective interventions, which are implemented in complementary ways at reasonable cost.

Between 2001 and 2008 the US Centers for Disease Control and Prevention's *Controlling Asthma in American Cities Project* (CAACP) sought to use a public health approach to reduce the burden of asthma in seven cities in the United States. Coalitions of diverse stakeholders at each site were funded and charged with developing and implementing a comprehensive set of asthma interventions based on local needs. Details of the project's purpose, design, strategic plan, and evaluation are described elsewhere (Herman 2011).

This chapter draws on the CAACP experience and the available literature to introduce and articulate questions implicit when planning a public health approach to impact asthma in a population. The chapter is organized into three main topics: Measures and Outcomes, Selection of Interventions, and Targeting Resources. Each topic has two or three subtopics that represent many of the key questions planners must address due to their importance to the planning process and the relative lack of guidance in the existing literature. The chapter does not seek to definitively answer or resolve these questions. Rather, its purpose is to explore the many issues and challenges that must be considered when planning an asthma intervention in a large group or population. The information presented is intended to help inform government agencies, non-governmental organizations, managed care organizations, health systems, community coalitions, and funders who are stakeholders in combating asthma, and also serves to identify areas for future research.

2. Measures and outcomes

2.1 What population-level outcome measures are available and appropriate for evaluating asthma interventions?

Small scale research trials typically form the basis for testing the efficacy of interventions among a relatively small group of individuals. In such settings tools that have been validated for clinical use, such as the Asthma Control Test (Liu et al., 2007), may be used to evaluate the intervention's impact. However, data from such clinical tools are seldom available for large populations of persons with asthma. Demonstrating an intervention's impact in a population requires data representative of the entire population at multiple points in time. As the development of population-level databases is typically expensive and time consuming, most disease-specific public health efforts use existing data available for general use. These might include (a) publicly available databases, such as vital statistics and disease registries, hospitalization discharge records, or emergency department data; (b) proprietary medical records or claims data; or (c) surveys from a representative sample of the population. The asthma-related data available for a given population might be quite limited. To ensure that a program's targeted outcomes are measurable, realistic, and consistent with the level of funding,available databases and the information needed from them should be identified at the beginning of the planning process.

Types of population-level measures of asthma can be conceptualized in a pyramid, with the least common events at the top (Fig 1) (Centers for Disease Control and Prevention 2001). Asthma mortality is seldom used as an outcome measure for small to moderate sized populations because it is too rare an event to detect statistically significant changes over time. Although also fairly rare, and representative of only the most acute cases of asthma, asthma hospitalizations (overnight stays in a hospital) are much more commonly used to measure success of an asthma intervention. In the U.S., hospitalization data are collected at the state level, although quality and availability varies (Love et al., 2008). Asthma related emergency department visits are 3–4 times as common as asthma related hospitalizations (Moorman et al., 2007). Two of the seven CAACP sites were able to obtain emergency department data from a central source. Other sites had to collect and collate information from individual hospitals and health plans serving their populations to estimate population ED utilization rates.

One site analyzed patterns in asthma medication prescription fills and estimated use in a population by utilizing a pharmacy chain's administrative database (Davis et al., 2011). Measures of asthma control, adherence to clinical guidelines, and types of symptoms and their frequency are available from national and a few state-level surveys, as well as some health plans, but are generally not available for a small geographic catchment areas (Centers for Disease Control and Prevention 2010c). Measures of underlying asthma severity, as it is defined by the NHLBI, can vary over time in an individual and would not be expected to necessarily change as the result of a clinical or public health intervention.

Several sites were able to monitor asthma-related school absences at the school-district level. Measures of asthma prevalence such as "lifetime prevalence" or "current asthma" change very slowly over time. Furthermore, since there is no known cure for asthma, nearly all of the available interventions focus on secondary prevention, rather than the primary prevention or elimination of the disease. Although the reduction of certain occupational exposures is now known to reduce risk of developing asthma, these exposures are limited and affect relatively few individuals.

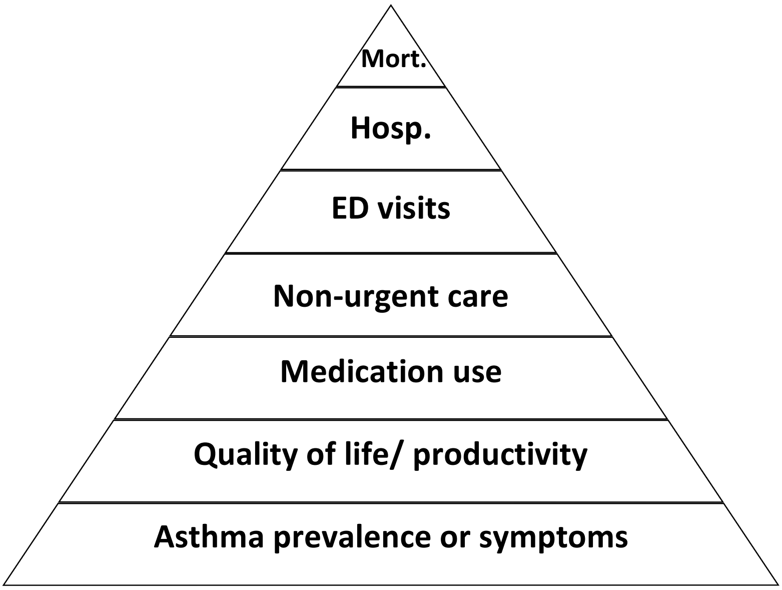


Fig. 1. Types of population level outcome measures for asthma

In some cases, seemingly "successful" asthma interventions may cause some outcome indicators to shift in an unintended direction. For example, providing asthma education to individuals who were previously unaware of asthma's seriousness or unfamiliar with its symptoms may lead to greater emergency department utilization or symptom reporting. A parent of a toddler with asthma, upon learning for the first time that asthma can be fatal, or about what happens to the airways during an asthma episode, may be more inclined to take her pre-verbal child to the emergency department upon hearing a wheeze.

Furthermore, the desirable direction for an indicator can be ambiguous. For example, most asthma interventions have aimed to increase use of inhaled corticosteroids (to prevent asthma episodes) and to decrease bronchodilator use (to relieve existing symptoms). Since bronchodilators are taken by individuals to relieve asthma symptoms, they are used as a marker of poorly controlled asthma. However, some individuals, prior to an intervention, do not have access to bronchodilators or are not using them prior to exercise or when they actually need them. For those individuals, filling a prescription for a bronchodilator and using one may be a positive change. Likewise, increased use of inhaled corticosteroids could mean better adherence to the asthma management plan by the patient or better adherence to clinical guidelines by the provider, but it may also mean a larger percentage of a population is developing persistent, as opposed to intermittent, asthma.

Outpatient visits as a measure of an asthma intervention's effectiveness can be ambiguous to interpret as well. One hopes to reduce urgent, unscheduled asthma visits and increase routine "well" visits, but many data systems do not make a distinction between the two types of visit. The reduction of emergency department visits is a common measure of success for asthma interventions but this reduction may also lead to an increase in non-emergent outpatient services as people shift to more appropriate preventive care. Although this shift should be seen as a positive behavioral change, it could appear to be "harmful" or costly if not viewed in the proper context.

Similarly, most patients hospitalized for asthma present at the emergency department. When someone is hospitalized for asthma, even if the patient initially presents at the emergency department, the hospital records this as an inpatient visit and not an emergency department visit. Thus, if a smaller percentage of patients are presenting to the emergency department with symptoms necessitating hospitalization it could actually work to increase emergency department visits, and vice versa.

Reduction of the cost of asthma is a derivative measure that may be paramount for some stakeholders.

Successful asthma interventions will decrease some preventable costs (e.g. acute care visits), but may lead to increased costs elsewhere (e.g. increased use of inhaled corticosteroids, preventative office visits). Public institutions and coalitions should be aware that stakeholders may not all agree on how to target costs. Cost shifting hospitalizations may cause acrimony between two otherwise well-intentioned organizations. For example, an asthma intervention that reduces asthma-related hospitalizations may help a health plan's bottom line, but may hurt the hospital's bottom line depending on the characteristics of the particular patient (private or public insurance, capitated or not capitated). Ultimately, the "best" measures to evaluate an intervention will depend on what is available as well as what the particular goals of the stakeholders are.

2.2 Scaling interventions: How many individuals need to be reached and with what effectiveness to achieve desired outcomes for various population sizes?

To produce change at the population level, the number of individuals reached and the effectiveness of the intervention(s) must be commensurate with population size. Even the most effective interventions will not result in population-level change unless enough people are affected. For example, even if a hypothetical intervention that is 100% effective at preventing a subsequent asthma hospitalization for one year were given to every patient hospitalized for asthmain a large city in an entire year, it would reduce hospitalizations the following year by only a small amount (Holgate 1999), as only a fraction of hospital admissions for asthma in a population in a given year are readmissions (Centers for Disease Control and Prevention 1997).Estimating the number of individuals who must be reached, and with what effectiveness, to demonstrate significant change in a particular population-level outcome is an important step in planning a community-based intervention.

This task differs from power calculations used to determine the minimal sample size or experimental group necessary to achieve statistical significance in a study. That calculation requires estimates of the population parameters. The intent here is to achieve an actual change in the population parameter itself. Estimating the number of individuals who must be reached to achieve that change requires the following three pieces of data: the prevalence of asthma in the population; the frequency of the event to be measured among those with asthma (e.g., the rate of hospitalizations, emergency visits, or office visits); and the random variation associated with that event. The change in the number of events that must be achieved in a given population (in order to exceed random variation and thus reach significance) can be calculated. The intended direction of change can be either positive or negative. The number of people who must be reached to achieve that change can be calculated from the change in the number of events needed, the frequency of that event, and the effectiveness of the intervention or project.

For example, recent national data indicate that the prevalence of asthma is 7.3% and approximately 66 office visits, 8.8 emergency department visits, and 2.5 hospitalizations occur per 100 persons with current asthma (Moorman et al., 2007). As shown in Table 1, in a hypothetical population of 500,000 (the approximate size of CAACP target populations), one would expect 36,500 people with current asthma (500,000 X 0.073); 24,090 office visits (36.500 X 0.66); 3,212 emergency department visits (36,500 X 0.088); and 913 hospitalizations (36,500 X 0.025) for asthma annually. Estimating the change in the number of events needed to achieve significance requires multiplying the standard error of that event by 1.96.[#] In Table 1, the standard error is approximated by taking the square root of the number of events in the population. This approximation, based on the Poisson distribution, is derived from the formula for relative standard error used for mortality data (Arias et al, 2003). Thus, in a population of 500,000, the number of emergency department visits that must be eliminated to reach significance ($p < .05$) is the square root of 3,212 multiplied by 1.96, or 111 events.

[#] 1.96 is the critical value of the z distribution for a two-tailed test at the 0.05 level of significance.

Population Size	Number with asthma†	Office visits for asthma‡		ED visits for asthma§		Hospitalizations for asthma¶	
		number	change**	number	change	number	change
5,000,000	365,000	240,900	962	32,120	351	9,125	187
2,500,000	182,500	120,450	680	16,060	248	4,563	132
1,000,000	73,000	48,180	430	6,424	157	1,825	84
500,000	36,500	24,090	304	3,212	111	913	59
250,000	18,250	12,045	215	1,606	79	456	42
100,000	7,300	4,818	136	642	50	183	26
50,000	3,650	2,409	96	321	35	91	19

† Based on a prevalence of 7.3%
‡ Based on a rate of 66 per 100 with asthma
§ Based on a rate of 8.8 per 100 with asthma
¶ Based on a rate of 2.5 per 100 with asthma
** Change = 1.96 · Standard Error; Standard Error ~ Square Root (number of expected events)

Table 1. Expected and associated change in the number of events required to reach significance in populations of various sizes*

The number of individuals who must be reached to achieve the necessary change in number of events can be estimated by dividing the needed change in number of events by the rate for the event. Table 2 illustrates this calculation for a population of 500,000. As stated above, to significantly reduce the population-based emergency department visit rate for asthma, at least 111 visits must be eliminated by the intervention(s). If the intervention is 100% effective in eliminating emergency department visits among those with asthma and there are 8.8 emergency department visits for every 100 with current asthma, then approximately 111 / 0.088 or 1,261 participants with current asthma must be reached. If less than 100% effective, that number is divided by the estimated effectiveness. An intervention that is 50% effective will need twice as many participants. Continuing with the above example, if the intervention eliminates half the emergency department visits among those enrolled, then 2,522 must be reached (1,261 / 0.50 = 2,522) to significantly decrease the emergency department visit rate for asthma. Further adjustments may be needed to allow for dropouts and incomplete participation, for population mobility (people who received the intervention moving out of the area, people who did not receive it moving in) as well as the uneven distribution of events in the population (some people having frequent events, others having none). Furthermore, directing interventions to populations most likely to experience the adverse event the intervention seeks to reduce would serve to decrease the total number of individuals who need to be reached.

---

*Refers to count of events such as a hospitalization, ED visit, or office visit

	Number of participants with asthma needed to demonstrate significant change in the event of interest*		
Effectiveness of the intervention	Office visits for asthma	ED visits for asthma	Hospitalizations for asthma
100 %	461 (304/0.66)	1,262 (111/ 0.088)	2,360 (59/0.025)
50 %	921	2,522	4,720
25 %	1,842	5,045	9,440

* Number needed = (number needed to change / rate of event) / effectiveness, e.g., for an intervention 50% effective in reducing emergency department (ED) visits for asthma, number needed = (111/0.088)/0.50 = 2522

Table 2. Number of participants with asthma needed for a population of 500,000
By effectiveness of the intervention (assuming 7.3% prevalence of asthma)

These calculations, although rough estimates, permit planners to determine whether the fit between target population size, available resources, and the outcome to be measured is realistic.

2.3 What external contextual factors are likely to influence the effectiveness of the intervention(s)?

A variety of external factors and changes over time can influence asthma-related outcomes. These factors may include but are not limited to demographic and economic changes, revision of reimbursement or coding policies, changes in Medicaid eligibility requirements, closing or opening of safety-net health service providers, concurrent interventions, changes in environmental exposures, and fluctuation in the intensity of cold/flu seasons (Johnston et al., 1996). Changes in a community that improve access to quality medical care or reduce environmental asthma triggers, for example, may make it difficult to attribute improved asthma outcomes to a project. Conversely, changes in external factors that negatively impact persons with asthma may mask the accomplishments of an asthma intervention.

Population movement out of a project area, a common occurrence, can theoretically affect outcome measures in a variety of ways. Each year 14% of people in the U.S. change their address with the rate tending to be higher in lower socioeconomic neighborhoods (U.S. Census Bureau 2010). Although individuals and families participating in asthma interventions may continue to benefit after leaving a project area, their improved outcomes would not be reflected in a population-based measure. Because asthma prevalence varies among ethnic groups (Davis et al., 2006), a change in racial or ethnic distribution may result in different rates of healthcare utilization that mirror a demographic shift rather than project impact. While the direction of change in asthma outcomes due to demographic changes is difficult to predict, increasing the estimated “number needed to reach” to accommodate the potential loss to follow up of individuals and families would be a conservative approach.

Over the seven-year CAACP project period, a number of external factors that had the potential to affect hospitalization rates and other outcome measures occurred in the CAACP sites. For example, new management at the major hospital at one of the sites relaxed the threshold for hospitalizing an asthma patient in order to fill more beds, effectively increasing the number of asthma hospitalizations quite significantly. When interpreted out of context, one could reasonably conclude that asthma in that community was getting worse and that the various interventions at that site were at best, not effective, and at worst, harmful. The sites addressed the complexity of external factors in a variety of ways. Davis et al. analyzed pharmacy-fill data using multiple complementary techniques (time trends, comparison of the project area with similar areas in the city, analysis by age group) to provide different perspectives and strengthen the attribution of improved patterns to the project (Centers for Disease Control and Prevention 2001). In the final analysis of hospitalization data, all sites will superimpose a contextual analysis of significant changes over the timeline of the projects' interventions and trends in hospitalization data. Although no analytic techniques can completely control for contextual factors, documenting and acknowledging them can facilitate a realistic interpretation of outcome data.

3. Selection of interventions

3.1 What criteria should be considered when choosing interventions to achieve population-level outcomes?

Reaching large numbers of people, in a variety of settings, in complementary and synergistic ways, and at reasonable cost, requires interventions at multiple levels(National Heart Lung and Blood Institute 2007). The Spectrum of Prevention (Box 1) (Cohen & Swift 1999) is one framework for categorizing the levels and types of interventions for asthma. Selecting interventions that are most likely to be effective is a critical part of the local planning process. Asthma clinical guidelines, review papers, and meta-analyses provide an overview of the evidence base behind different types of interventions, and are based on a systematic review of multiple studies. These publications typically give greater weight to randomized, controlled trials, and to studies with large sample sizes. Planners should also base their assessment of effectiveness on individual papers and when possible, discussions with authors and program staff of the intervention(s) being considered. When doing so, a critical assessment of the methodology and data collection methods of published studies is important.

For example, many published evaluations of asthma interventions rely on self-reported behaviors or symptoms. The accuracy and precision of self-report is likely to vary by the type of question, the person collecting the information, time transpired since the reported event, and provider of the information (e.g., a patient or guardian) (Mathiowetz & Dipko 2000). Evaluations that do not have a control or comparison group might show a positive change based on a phenomenon called "regression to the mean," meaning that, when a series of events is tracked, the events will tend to return to a predictable mean on their own even without intervention (Tinkelman & Wilson 2004). Regression to the mean can lead to an incorrect conclusion that attributes an asthma outcome to an intervention when it was actually due to chance. This effect might be especially pronounced in asthma interventions that focus on individuals with a recent event, or with high utilization at baseline (Tinkelman & Wilson 2004).For example, if a given patient is recruited into an intervention as a result of a recent asthma hospitalization, statistically it is unlikely that this same patient would have re-experienced another asthma hospitalization during or after the study period. Hospitalizations due to asthma are relatively uncommon, even among those with previous asthma hospitalizations.

- **Influencing Policy and Legislation**
  - Clean outdoor air policies and legislation
  - Clean indoor air policies and legislation
  - Occupational regulations
  - Healthcare delivery and financing
- **Strengthening Social/Organizational Practices**
  - Quality improvement initiatives
  - Information systems
  - Local school policies
- **Educating/Training Healthcare Providers**
  - Educating primary care providers
  - Specialized training for medical residents
  - Educating pharmacists
  - In-services for school and childcare center staff
- **Promoting Community Education**
  - Group asthma education in community settings
  - Group asthma education in schools
  - Social marketing, public service announcements
- **Strengthening Individual Skills and Knowledge**
  - Case management (home, school, clinic)
  - Patient education in hospitals and emergency rooms
  - Phone follow up or nurse monitoring
- **Clinical Care/Treatment**
  - Specialty asthma or allergy clinics, mobile clinics
  - Telehealth applications

Box 1. Types of asthma interventions using the Spectrum of Prevention as a framework to categorize(Cohen & Swift, 1999).

The effectiveness of any intervention is dependent upon context (Wang et al., 2006). In their comprehensive review of community based public health interventions, Sorensen, Emmons, and Dobson present a persuasive argument that the efficacy-based research paradigm that dominates in research journals may not be the most appropriate way to evaluate public health interventions (Sorenson et al., 1998), and tend to produce interventions that are intensive and expensive. Efficacious interventions conducted under rigorous study design with carefully screened and motivated participants, by the most skillful professionals, may prove ineffective in other settings (Glasgow et al., 1999; Sorenson et al., 1998; Starfield 1998). Even if they are highly effective, they will have little population based impact if they cannot be widely adopted (Glasgow et al., 1999).Planners should give greater weight to interventions that have proven effective in similar environments and circumstances and where applicable and when possible, pilot test them on the population the planners hope to impact.

Local political considerations may prove be more influential in the final selection of the intervention than the literature. The CAACP sites selected their interventions during a planning process that involved reviewing the literature on asthma interventions, conducting a needs assessment, soliciting stakeholder interests, and performing small pilot studies of proposed interventions. Some coalitions reached agreement on the mix of interventions through consensus; others followed structured procedures. As documented in annual reports, the Minneapolis/St Paul site actively engaged 115 people in intervention selection. Six workgroups met monthly for seven months to move through a formal process that resulted in a prioritized list of interventions, an evaluation plan, and a proposed budget. A leadership team then identified areas of overlap and synergy from the six plans to compose a strategic project plan. All sites had a coalition-based process that balanced local needs and resources, stakeholder preferences, and evidence supporting intervention effectiveness.

While most interventions were evidence-based, CAACP sites implemented some interventions for which an evidence base was not yet available. They justified these interventions with logic models that linked the interventions and desired outcomes (Cheadle et al., 2003). For example, the St. Louis site's needs assessment identified healthcare system fragmentation as one of the target area's major challenges. The site implemented the Asthma Friendly Pharmacy intervention, although not supported by pre-existing evidence, because it appeared to be a critical link in a communications network among patients and their families, healthcare providers, and schools. Evaluation of that intervention focused on its success in establishing and maintaining those lines of communication (Berry et al., 2011) and now contributes to the evidence base for pharmacy interventions.

3.2 How should resources be allocated across multiple asthma interventions?

Little practical guidance exists for allocating public health resources for asthma across the range of interventions listed in Box 1, and no studies have tested the effectiveness of different combinations of interventions. Given the number of possible combinations of interventions, differences in communities, timing of interventions, and variations in external factors affecting outcomes, an empirical answer to the question of the most effective mix of interventions is unlikely.

The CAACP projects chose interventions that addressed different levels of the Spectrum of Prevention. All sites provided individualized family and home asthma services that focused on asthma self-management training and indoor-trigger reduction. The Philadelphia, Richmond, and Minneapolis/St.Paul sites reached out to parents and community members by providing asthma classes in community settings; Philadelphia site staff made extensive use of local communication networks (radio, newsletters, newspapers) to increase community awareness of asthma and the CAACP. Training for primary healthcare providers on NAEPP guideline implementation was also included in all strategic plans. The Oakland and Minneapolis/St Paul sites implemented interventions to institutionalize the reinforcement of key asthma messages into routine care for hospitalized and emergency department patients respectively, and the Northern Manhattan, Minneapolis/St Paul, and Oakland sites succeeded in institutionalizing asthma-friendly policies and procedures in their respective school systems. The Chicago site worked at the policy level to support smoking bans and the regulation of power plants and demolition sites.

CAACP sites noted an implicit tradeoff between designating comprehensive resources to a smaller number of people (i.e., high intensity) or fewer resources per person to a greater number of people (i.e., high reach). High-intensity interventions included, for example, individualized home-based medical and social support for families, and practice-based systems-change interventions for healthcare providers. High-reach interventions included group trainings and classes, policy-based interventions, and mass media.

As noted by Glasgow, the "high intensity" attributes that help to make interventions efficacious in a research setting—time, expertise, resources, commitment-- may actually work *against* the likelihood that they will be effective in less ideal settings (Glasgow et al., 1999). In contrast, low intensity interventions that can be delivered to large numbers of people may have a more pervasive public health impact (Hatziandrew et al., 1995; Vogt et al., 1998). Conceptually, it is helpful to define the impact (I) of an intervention as the product of an intervention's efficacy (E) and its reach (R) (the percent of the population receiving an intervention): I=R*E (Abrams et al., 1996). Even within the context of one type of intervention there are considerations about reach and intensity. For example, among CAACP sites, the number of hours dedicated to group training of medical providers ranged from 1–8 hours, and length of home-based support for families ranged from 1–18 visits. Intensive interventions, although generally effective for participating individuals or families, cost more per person and may have had a limited effect on population measures when the number of people or clinics reached is a small proportion of the population (Glasgow et al., 1999). Low-intensity interventions may not be effective in changing behaviors or achieving outcomes, or may take many years to demonstrate an effect (Glasgow et al., 1991).

Glasgow's RE-AIM framework expanded on the I=R*E equation by adding three additional dimensions that more fully capture the real-world impact of an intervention. The additional dimensions are: adoption ("A", the percentage of organizational settings that will adopt an intervention), implementation ("I", the extent to which the intervention is implemented as intended in the real world), and maintenance ("M", the extent to which an intervention is sustained over time) (Glasgow et al., 1999). Traditional intervention trials emphasize only the efficacy component at the exclusion of the other components. Although efficacy may be of most interest to clinicians, it is insufficient to evaluate the impact of an intervention in a population. The RE-AIM framework provides a conceptual public health model for determining what programs are worth sustained investment in the real world.

4. Targeting and reaching individuals and communities

4.1 What factors should be considered when deciding which individuals or groups of individuals to target within a given community or project area?

Asthma is a complex and heterogeneous condition with multiple phenotypes. It can be described in terms of its control and severity, types of symptoms, frequency and intensity of exacerbations, impairment, and responsiveness to medications. Furthermore, these characteristics change over time for every individual with asthma. Interventions that provide a service must decide which individuals with asthma should receive the service in order to maximize the intervention's goals. In a study of limited scale, where the goal is to maximize efficacy, there is an incentive to recruit individuals that are most likely to respond favorably to the intervention. On the other hand, if the goal is to impact asthma outcomes for a population, then identifying and targeting individuals who are most at-risk for the outcome that the intervention is designed to improve is essential for maximizing the intervention's impact.Even small differences in characteristics between participants and non-participants can have a significant impact on an intervention's effectiveness when conducted on a large scale (Vogt et al., 1998).

The national asthma guidelines frequently refer to "high risk" individuals, but use the term inconsistently (National Heart Lung and Blood Institute 2007). One challenge planners face is the lack of correlation among different measures of asthma. Underlying severity and measures of current impairment, such as symptoms, functional limitations, or quality of life, are not reliable predictors of future risk of asthma exacerbation or adverse events (National Heart Lung and Blood Institute 2007). Furthermore, different adverse events may not necessarily correlate with one another; someone who frequently misses school because of asthma may not necessarily be at risk for a hospitalization.

Selecting an age group on which to focus is another consideration for maximizing the population impact of an intervention. Young children have the highest rates of asthma-related acute visits and have likely received the greatest amount of attention from government and private funders in the US, but this risk decreases significantly in the teen and adult years (Centers for Disease Control and Prevention, 1997). It is difficult to predict which children will be at "highest risk" as they age, and therefore which ones should be the targets of intervention. Provision of resources to a given child may thus have only a short-term impact. Senior individuals with asthma have the highest asthma-related mortality, and hospitalization costs that are on average two times that of young children per visit (Bahadoriet al., 2009)but for reasons that can only be speculated, are far less commonly the subjects of intervention trials.

Risk assessment models have used empirical data to successfully predict, in the short term, who is most at risk for various outcomes and thus most likely to have an impact on those outcomes in the future (Eisner et al., 2002; Li et al., 1995; Magid et al., 2004; Vollmer et al., 2002). These models include many variables, such as demographic data, various indicators of current asthma control, prescription patterns, and healthcare utilization history. Such models are limited because they differ for each population and require a richness of data that may not be available to the intervention planners.

Despite the fact that the CAACP sites had the same outcome goals, their definition of "high risk" and the selection criteria and methods that the different sites used for determining eligibility for interventions varied greatly. For instance, the criteria the different sites used for enrollment in a home-based asthma case management program included one or more of the following: asthma severity, various measures of asthma control, history of asthma-related hospitalizations, socioeconomic factors, school absences, and physician referral. The more resource-intensive interventions were generally more selective (CAACP personal communication 2008), in keeping with evidence that most asthma morbidity and cost can be attributed to relatively few individuals (Smith et al., 1997).

4.2 How can projects effectively reach large numbers of individuals to participate in interventions?

The efficient identification, recruitment, and retention of a large number of individuals who drive adverse outcomes are particularly important when the goal is to impact population outcomes. Although the challenges of participant recruitment and retention are not unique to public health projects, they can dramatically limit a project's ability to achieve a population-level impact. Individuals and groups who are high priority from a public health perspective may be the least able to participate in interventions.

Schools and daycare centers were logical partners for CAACP sites because they offered access to a large proportion of a community's children and had an interest in reducing absences caused by poorly controlled asthma. Three sites—Oakland, Northern Manhattan, and Chicago—used questionnaire-based case identification of students with asthma in the schools as the primary method of identifying and recruiting students for interventions. Those sites also used the frequency of self-reported symptoms as a means of prioritizing students for interventions of high intensity and cost. Oakland was successful at engaging 76% of all students identified with asthma. In the Northern Manhattan site, 35% of parents of children identified with asthma participated in at least one asthma education activity offered by the daycare center.

Hospitals are potential venues for identifying and educating large numbers of patients with uncontrolled asthma(National Heart Lung and Blood Institute 2007), as patients can be easily enrolled in an intervention while in the hospital or ED without the traditional barriers of time and transportation (Castro et al., 2003; Teach et al., 2006).The Oakland CAACP site found that working with patients who were hospitalized for asthma to be an effective way to reach high risk children with asthma.

Many health plans, have an interest in reducing costs and improving performance measures, and generally have data that can be used to identify high-risk members or evaluate interventions (Hoppin et al., 2007; Kantor 2007; United States Environmental Protection Agency 2008a, 2008b). Some health plans referred children to CAACP services.Other CAACP site recruitment methods included referrals by medical providers (Chicago), distribution of flyers (Philadelphia), partnerships with social service agencies (Richmond and St Louis), outreach to parents of children in daycare (Northern Manhattan and Richmond), and door-to-door solicitation (Minneapolis/St Paul).

4.3 What are some considerations for addressing population level asthma disparities?

The reduction of disparities in disease burden among socio-demographic groups is an important public health goal, and one that requires focusing resources on certain populations and communities. Although all CAACP sites were selected because their populations suffered a high, disparate burden of asthma, they received no direction about addressing disparities within the sites. Many of the children with asthma in those communities were considered to have a disparate burden of asthma because of low socioeconomic status (SES), unstable living situations, disorganized families, limited access to care, poor housing conditions, or some combination of these determinants of health.

Social, economic, and psychological factors are important determinants of whether an individual will benefit from an intervention (Weil et al., 1999). Individuals in challenging circumstances may have unique incentives or disincentives that affect behavior. For example, families enrolled in government insurance may have a low threshold for deciding to use the ED, which requires no appointment and whose cost is negligible to families, and may in effect use the ED as a replacement for regular primary and preventive care. This phenomenon may be a reason that so many patients present to the ED with mild symptoms (Macias et al., 2006).The economic dynamics which affect behavior likely vary by insurance provider and by extension, by state and country.

All CAACP sites offered referrals to address substance abuse, psychiatric, housing, and financial problems to the extent that those services were available locally, but varied in their efforts to actively address social determinants of health. The Richmond site was most proactive in partnering with a local agency to provide a range of social services as well as asthma education and case coordination for children "who had failed every other intervention." Similarly, the St Louis site contracted with a social service agency to address families' socioeconomic priorities to enable the families to also focus on controlling their children's asthma.

Reaching people and groups who experience health disparities primarily for social and economic reasons often requires labor-intensive, time-consuming, costly interventions. Intensive interventions focused on a small group of people may not necessarily be the most cost effective strategy if the goal is to improve asthma for a diverse population. However, if the priorities are reducing disparities and/or reaching individuals who are most "in need", then interventions that address recidivism and include social services are indicated. It is important to clarify among the stakeholders early in the planning stages if reduction of health disparities is the goal.

5. Conclusion

Public health efforts aiming to improve population-level outcomes are fundamentally different from projects seeking to improve outcomes for patients or for study participants, yet little guidance exists for making choices necessary for planning and implementing public health interventions addressing asthma. Consideration of appropriate population-based outcome measures, the number of people needed to reach to improve outcome measures, the choice and mix of interventions, and priorities in targeting and methods of recruiting participants all need to be addressed explicitly during the planning phase. The authors hope that future research will provide additional guidance on implementing projects aiming to reduce the burden of asthma at a population level.

6. References

Abrams DB, Orleans CT, Niaura RS, Goldstein MG, Prochaska JO, Velicer W. (1996). Integrating individual and public health perspectives for treatment of tobacco dependence under managed health care: A combined stepped care and matching model. *Ann Behav Med.* 18, 4, (Fall 1996), pp. 290-304.

Arias E, Anderson RN, Kung H, Murphy SL, Kochanek KD. (2003). Deaths: final data 2001. *National Vital Statistics Reports*. 52, 3, (September 2003), pp. 111.

Asher MI, Montefort S, Björkstén B, Lai CK, Strachan DP, Weiland SK, Williams H. (2006). Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys.*Lancet*. 368, 9537, (August 2006), pp. 733-743.

Bahadori K, Doyle-Waters MM, Marra C, Lynd L, Alasaly K, Swiston J, FitzGerald JM. (2009).Economic burden of asthma: a systematic review. *BMC Pulm Med*. 19, 9, (May 2009), pp. 24.

Berry T, Prosser T, Wilson K, Castro M. (2011). Asthma friendly pharmacies: A model to improve communication and collaboration among pharmacists, patients, and health care providers. *Journal of Urban Health*.88, 1, (February 2011), pp. 113-125.

CAACP site directors or investigators personal communication, February–September 2008.

Castro M, Zimmerman NA, Crocker S, Bradley J, Leven C, Schechtman KB. (2003). Asthma intervention program prevents readmissions in high healthcare utilizers. *Am J RespirCrit Care Med*. 168, 9, (November 2003), pp. 1095-1099.

Centers for Disease Control and Prevention. (1997). Asthma hospitalizations and readmissions among children and young adults—Wisconsin 1991-1995.*MMWR*. 46, 31, (August 1997), pp. 726-9.

Centers for Disease Control and Prevention. (2001). Adapted from "A Public Health Response to Asthma," PHTN Satellite Broadcast, Course Materials 2001.

Centers for Disease Control and Prevention. (2010a). Logic Model, Accessed on 25 Feb 2010, Available from: <http://www.cdc.gov/eval/resources.htm>

Centers for Disease Control and Prevention. (2010b). Guide for state health agencies in the development of asthma programs, Accessed on 25 Feb 2010, Available from: <http://www.cdc.gov/asthma/pdfs/asthma_guide.pdf>

Centers for Disease Control and Prevention. (2010c). Behavioral Risk Factor Surveillance System, Accessed on January 30, 2010, Available from <http://www.cdc.gov/asthma/survey/brfss.html>

Cheadle A, Beery W, Greenwald H, Nelson GD, Pearson D, Senter S. (2003). Evaluating the California Wellness Foundation's Health Improvement Initiative: a logic model approach. *Health PromotPract*. 4, 2, (April 2003), pp. 146-56.

Coffman JM, Cabana MD, Halpin HA, Yelin EH. (2008). Effects of asthma education on children's use of acute care services: meta-analysis, In: *Pediatrics* 121:575-586 [serial online March 4, 2008], Available from: <http://www.pediatrics.org/cgi/content/full/121/3/575>

Cohen L, Swift S. (1999). The spectrum of prevention: developing a comprehensive approach to injury prevention. *Inj Prev*. 5, 3, (September 1999), pp. 203-207.

Davis A, Kreutzer R, Lipsett M, King G, Shaikh N. (2006). Asthma prevalence in Hispanic and Asian American ethnic subgroups: results from the California Healthy Kids Survey. *Pediatrics*. 118, 2, (August 2006), pp. 363-370.

Davis SQ, Krishnan JA, Lee K, Persky V, Naureckas ET. (2011) Effect of a community-wide asthma intervention on appropriate use of inhaled corticosteroids. *Journal of Urban Health*.88, 1, (February 2011), pp. 144-155.

Eisner MD, Ackerson LM, Chi F, Kalkbrenner A, Buchner D, Mendoza G, Lieu T. (2002). Health-related quality of life and future health care utilization for asthma. *Ann Allergy Asthma Immunol*.89, 1, (July 2002), pp. 46-55.

Glasgow RE, Hollis JF, McRae SG, Lando HA, LaChance P. (1991). Providing an integrated program of low intensity tobacco cessation services in a health maintenance organization. *Health Educ Res*. 6, 1, (March 1991), pp. 87-99.

Glasgow R, Vogt T, Boles S. (1999). Evaluating the public health impact of health promotion interventions: The RE-AIM Framework. *American Journal of Public Health*. 89, 9, (September 1999), pp. 1322-1327.

Hatziandrew EJ, Sacks JJ, Brown R, Taylor WR, Rosenberg ML, Grahan JD. (1995). The cost effectiveness of three programs to increase use of bicycle helmets among children.*Public Health Rep*. 110, 3, (May-June 1995), pp.251-259.

Herman E. (2011). Conceptual framework of the Controlling Asthma in American Cities Project.*Journal of Urban Health*. 88, 1, (February 2011), pp. 7-15.

Holgate S. (1999). Difficult asthma: the unanswered questions. *Difficult Asthma*. London: Martin Dunitz Ltd. 1999.

Hoppin P, Jacobs M, Stillman L. (2007).Investing in Best Practices for Asthma: A Business Case for Education and Environmental Interventions. As*thma Regional Council of New England Health Resources in Action*. (July 2007), pp. 1-21.

Institute of Medicine. (1988). *The Future of Public Health*. Washington, DC: National Academy Press; 1988.

Johnston SL, Pattemore PK, Sanderson G, Smith S, Campbell MJ, Josephs LK, Cunningham A, Robinson BS, Myint SH, Ward ME, Tyrrell DA, Holgate ST. (1996) The relationship between upper respiratory infections and hospital admissions for asthma: a time-trend analysis. *Am J RespirCrit Care Med*. 154, (September 1996), pp. 654–660.

Kantor A. (2007). Breathing easier. *AHIP Cover*. 48, 3, (May-June 2007), pp. 50-2,54,56.

Li D, German D, Lulla S, Thomas RG, Wilson SR. (1995). Prospective study of hospitalization for asthma: a preliminary risk factor model. *Am J RespirCrit Care Med*. 151, (March 1995), pp. 647-55.

Liu AH, Zeiger R, Sorkness C, Mahr T, Ostrom N, Burgess S, Rosenzweig JC, Manjunath R. (2007). Development and cross-sectional validation of the Childhood Asthma Control Test. *J Allergy ClinImmunol*. 119, 4, (April 2007), pp. 817-25.

Love D, Rudolph B, Shah G. (2008). Lessons learned in using hospital discharge data for state and national public health surveillance: Implications for Centers for Disease Control and Prevention Tracking Program. *J Pub Health Management Practice*. 14, 6, (Nov-Dec 2008), pp. 533-542.

Macias CG, Caviness AC, Sockrider M, Brooks E, Kronfol R, Bartholomew LK, Abramson S, Shearer W. (2006). The effect of acute and chronic asthma severity on pediatric emergency department utilization.*Pediatrics*. 117, 4, (April 2006), pp. S86-95

Magid DJ, Houry D, Ellis J, Lyons E, Rumsfeld JS. (2004). Health-related quality of life predicts emergency department utilization for patients with asthma. *Ann Emerg Med*. 43, 5, (May 2004), pp. 551-557.

Mathiowetz NA, Dipko SM. (2000). A comparison of response error by adolescents and adults: findings from a health care study. *Med Care*. 38, 4, (April 2000), pp. 374-382.

McNairm J, Ramos C, Portnoy J. (2007). Outcome measures for asthma disease management.*CurrOpin Allergy ClinImmunol*. 7, (June 2007), pp. 231-235.

Moorman JE, Rudd RA, Johnson CA, King M, Minor P, Bailey C, Scalia MR, Akinbami LJ; Centers for Disease Control and Prevention (CDC). (2007). National surveillance for asthma—United States,1980-2004. *MMWR SurveillSumm*. 56, 8, (October 2007), pp.1-54.

National Heart Lung and Blood Institute. Expert Panel. *Report 3 (EPR 3): Guidelines for the Diagnosis and Management of Asthma*. Bethesda, MD: National Institutes of Health; 2007.

Smith DH, Malone DC, Lawson KA, Okamoto LJ, Battista C, Saunders WB. (1997). A national estimate of the economic costs of asthma.*Am J RespirCrit Care Med*. 156, (September 1997), pp. 787-793.

Sorensen G, Emmons KM, Dobson AJ. (1998). The implications of the results of community intervention trials.*Annu Rev Public Health*.19, (1998), pp.379-416.

Starfield B. (1998). Quality of care research: internal elegance and external relevance [commentary]. *JAMA*. 208, 11, (September 1998), pp. 1006-1008.

Teach SJ, Crain EF, Quint DM, Hylan MI, Joseph JG. (2006). Improved asthma outcomes in a high-morbidity pediatric population: results of an emergency department-based randomized clinical trial. *Arch PediatrAdolesc Med*. 160, 5, (May 2006), pp. 535-541.

Tinkelman D, Wilson S. (2004). Asthma disease management: regression to the mean or better? *American J Mang Care*. 10, 12, (December 2004), pp. 948-54.

U.S Census Bureau. (2010). Geographic Mobility Between 2004 and 2005, In: *Population Profile of the United States: Dynamic Version*, Accessed on 25 Feb 2010, Available from <http://www.census.gov/population/www/pop-profile/files/dynamic/Mobility.pdf>

United States Environmental Protection Agency. (2008a). *2008 National Asthma Forum Change Package: The System for Delivering High Quality Asthma Care*. Washington, DC: United Sates Environmental Protection Agency; 2008.

United States Environmental Protection Agency. (2008b) *Communities in Action for Asthma-Friendly Environments: A Systems-Based Approach for Creating and Sustaining Effective Community Asthma Programs*. Washington, DC: United States Environmental Protection Agency.

Vogt TM, Hollis JF, Lichtenstein E, Stevens VJ, Glasgow RE, Whitlock E. (1998). The medical care system and prevention: The need for a new paradigm. *HMO Practice*. 12, 1, (March 1998), pp. 5-13.

Vollmer WM, Markson LE, O'Connor E, Frazier EA, Berger M, Buist AS. (2002). Association of asthma control with health care utilization: a prospective evaluation. *Am J RespirCrit Care Med*. 165, 2, (January 2002), pp. 195-199.

Wang S, Moss JR, Hiller JE. (2006). Applicability and transferability of interventions in evidence-based public health.*Health Promot Int*. 21, 1, (March 2006), pp. 76-83.

Weil CM, Wade SL, Bauman LJ, Lynn H, Mitchell H, Lavigne J. (1999).The relationship between psychosocial factors and asthma morbidity in inner-city children with asthma.*Pediatrics*. 104, 6, (December 1999), pp. 1274-80.

An Integrated Theoretical Framework to Describe Human Trafficking of Young Women and Girls for Involuntary Prostitution

Thozama Mandisa Lutya[1] and Mark Lanier[2]
[1]*Department of Social Work and Criminology, University of Pretoria, Pretoria*
[2]*Department of Criminal Justice, College of Arts & Sciences, University of Alabama, AL*
[1]*South Africa*
[2]*USA*

1. Introduction

Human trafficking permeates diverse institutions whose systematic operations are entwined into a multitude of activities. A combination of theories should provide an integrated explanation of the occurrence of human trafficking. Although many forms of trafficking exist, we focus primarily on trafficking of women for involuntary prostitution. Bruckett and Parent (2002:7) are of the opinion that apart from the description of the processes, practices, and routes of human trafficking there has been a lack of consistency regarding the theoretical framework for understanding human trafficking. Conventional theory and methods suggest that strategies to conduct research on forced migrants require multi-disciplinary and interdisciplinary approaches, which at times may be divergent (Van Impe, 2000:124). The factors that enable human trafficking to occur vary and are interdependent and interconnected (Stop Violence Against Women (SVAW), 2008:1; Truong, 2001:34-35; Van Impe, 2000:117-118). It is possible that human traffickers observe trade in human beings as a profitable area to generate income, especially when considering that few human traffickers are arrested, prosecuted and sentenced for this crime. Ineffective criminal justice and community response to human trafficking strengthens the trafficking process, increases abuse of trafficked persons and allows human traffickers to generate financial proceeds from the crime. An integrated model to explain human trafficking appears to be a logical step towards an understanding of the crime. Current research explanations of the process of human trafficking are often informed by individual researchers' own theoretical framework creating an impression of human trafficking as a single-dimensional type of crime. This paper will provide a broad integrated framework which considers the stages at which human trafficking for involuntary prostitution occurs. An integrated framework should help enhance the prevention and control strategies utilized to reduce human trafficking of women for involuntary prostitution.

2. Definition of concepts

2.1 Human trafficking

Article 3 of the UN Palermo Protocol (2002:2), defines human trafficking in persons to mean: the recruitment, transportation, transfer, harboring or receipt of persons, by means of the threat or use of force or other forms of coercion, of abduction, of fraud, of deception, of the abuse of power or of a position of vulnerability or of the giving or receiving of payments or benefits to achieve the consent of a person having control over another person, for the purpose of exploitation. Consent is not necessary where exploitation, fraud, deception, and abuse of vulnerability have been involved. The various stages at which human trafficking for involuntary prostitution occurs are followed to build a multi-theoretical approach to human trafficking of young women and girls for involuntary prostitution.

2.2 Sexual exploitation

Article 3 of the UN Palermo Protocol (2002:2) defines sexual exploitation to include, at a minimum, the exploitation of the prostitution of others or other forms of sexual exploitation such as forced marriages, sexual slavery or servitude and mail order brides. For the purpose of this paper, reference is only made to women and girls sexual exploited through involuntary prostitution. Therefore prostitution, which is a voluntary sex work performance, is not equated with sexual exploitation. This paper also recognizes that prostitutes could be trafficked but is concerned with young women and girls forced into prostitution by human traffickers.

2.3 Theory intergration

The process theory integration entails merging of concepts from diverse disciplines and theories to explain a crime which involves a high contingent of perpetrators. Theory integration is a process of combining the best elements of existing theories to better explain the causes of criminal behaviour (Brown, Esbensen & Geis 2007:410; Lanier & Henry 2009: 382). This paper has merged concepts from, rational choice, victimology, demand theory, constitutive criminology and economic theories to explain human trafficking of young women and girls for involuntary prostitution. Human trafficking for involuntary prostitution requires a network of variables in order to construct a comprehensive view of its occurrence.

2.4 Women and girls

The concept woman refers to a person of female gender above the age of 18 and a girl refers to a female child under the age of 18. In this paper, the category of women described could range between the ages of 18 and 24 and girls refer to children ranging between the ages of 10 to 18. South African research on human trafficking of women and girls has revealed that children as young as 10 are trafficked. Although young women of all ages and nationalities could be trafficked it is pointed by ILO (2005) that the demand for prostitutes prefers girls younger than 24 years.

3. Theoretical framework

From the theories explicated below an integrated framework will be created to explain human trafficking of young women and girls for involuntary prostitution. Concepts from, rational choice, victim vulnerability, economic theory and constitutive theory will be selected to describe the sequences of events followed during the commitment of human trafficking for involuntary prostitution. It is pointed out by Lanier and Henry, (2004:343) that when a crime is an outcome of several different causes, an integrated framework or a conceptual absorption approach is required to analyse the sequential chain of events. In this regard, Lanier and Henry (2004) explain that the purpose of integrating theories is to present an interaction of probabilities from different theoretical perspectives that could explain the factors contributing to a person committing a crime. Eventually a recommendation of an epi-criminological standpoint as a strategy to respond to human trafficking of women for involuntary prostitution is made.

3.1 Rational choice theory

Rational choice theories postulate that criminals are rational beings who make decisions to commit crime based on the costs and benefits involved in the process of crime perpetration. Deterministic in nature, criminal decision making process is based on free will, which necessitates observation of opportunities, circumstances and situations that could affect the successful perpetration of the planned crime, (Lanier & Henry, 2004:90). It is pointed out by Brown, Esbensen and Geis, (2008:213) that rational decision making pertaining to crime also involves the choice of the victims determined by the type of crime, modus operandi, where and when to commit it and what to do afterwards. That means the criminals may first observe the accessibility to potential victims, location, the time at which they are at most vulnerable, the appropriate method that could provide entry with ease and how to safeguard their criminal activities from criminal justice authorities and other capable guardians. However, some rational theorists have argued that criminals differ in the choices they make based on their perceptions, motives, skills and abilities to read opportunities as situations guide their decisions making processes, (Lanier & Henry 2004:90). For the purpose of this paper: rational decision making, free will, cost and benefits are three variables that will help to build an integrated framework to explain human trafficking of women and girls for involuntary prostitution. The manner in which human traffickers select their victims is based on the gains they could get from the crime and vulnerability of potential victims. Nevertheless, there needs to be another theoretical explanation to describe the victim vulnerability of potential victims.

3.2 Demand theory

The demand for prostitutes can be classified into three categories: users or purchasers of sex, profiteers from selling sex, and socio-cultural attitudes towards sex, (Hughes 2004). Users or purchasers refers to persons who pay prostitutes to render a sexual service; brothel owners and pimps comprises of profiteers from selling sex and academics and media reporting and writing about prostitutes form part of socio-cultural attitudes towards sex, Hughes (2004). The motives behind purchasing or owning prostitutes and depicting sex services in the manner in which writers do, may contribute to human trafficking of young women and girls for involuntary prostitution. The purchasers of sex, cultural attitudes associated with prostitution, and violence towards women are three factors that this paper has identified to explain the increasing demand for prostitutes. It is deduced from these factors that a certain category of users of prostitutes do not necessarily separate caged prostitutes from voluntary prostitutes but could be more concerned with receiving sexual services from sex workers.

3.2.1 The purchasers of sex

The need for sex, cultural meaning associated with prostitution and violence towards trafficked victims are described in this paper as four distinguishing characteristics of users of prostitutes. Firstly, whatever reasons drawing men to prostitutes, Hughes, (2004:16) is of the opinion that the users of prostitutes are a heterogeneous group with different needs and motives towards sex with prostitutes. Of primary importance to users of prostitutes is the need to fulfil a personal inadequacy, need or desire. It is pointed out by Groom and Nandwani, (2006: 366); Hughes, (2004); Mansson, (2006:89) as well as Macleod, Farley, Anderson and Golding, (2008:14-18) that men's current intimate relationships experiences, desire for unfamiliar sex, acceptance of rape myths and sexual violence towards prostitutes, perceptions of prostitutes and prostitution, and lack of emotional connection could motivate men to solicit prostitutes. The need for prostitutes may contribute to human trafficking of young women and girls for involuntary prostitution. However, some users of prostitutes may go beyond the point of experimentation and sexual fulfilment by perpetrating acts of violence towards prostitutes such as: physical, emotional and sexual violence.

Secondly, there seems to be a cultural meaning associated with prostitution. The norms and values of men who purchase sex and the moral significance they attach to prostitution are an important contributing factor towards the demand for prostitution. It is pointed out by Coy et al, (2007: 19) as well as Macleod, et al. (2008) that the prostitute user's perception generally equates men, sexual aggression, and entitlement as cultural values defining superior manhood. According to the authors it is possible that the users of prostitute, who hold such beliefs adhere to rape myths, are mentally programmed to dominate women, have the desire to use prostitutes to revenge towards women who had wronged them in the past. Moreover, in societies where prostitutes are seen as a moral abomination they are least likely to receive sympathy for the violence they may experience at the hands of clients. Clients perceive prostitutes as morally different from other women - free spirited, fatally flawed and seem to sell their bodies in order to get money therefore deserve violence perpetrated towards them, (Macleod, et al. 2008:21).

Thirdly, despite the intention to fulfil personal inadequacies that cannot be satisfied in a normal relationship, prostitutes seem to experience physical, emotional and sexual violence from clients. However, the gendered violence endured by trafficked victims at the hands of users should be seen differently from intimate partner violence. Within this transactional context, the victims and the perpetrators are strangers. A short term agreement is the basis of their interaction. However, similarities between intimate partner violence and violence against trafficked victims are defined by the fact that men are known to use violence against women as a strategy to reassert authority weakened by their daily experiences. On the other hand the experience of victimisation from a victim's perspective is worsened by the type of work she is forced to perform, user's violence and the violence encountered from human traffickers. Eventually, women's bodies as objects to fulfil a man's desire for sex and proneness of prostitutes to client violence are essential elements describing the vulnerability of trafficked young women and girls.

Notwithstanding the cultural significance attached to prostitution and the moral abomination of prostitutes, the question that requires empirical scrutiny is whether users would refrain from purchasing sex if they knew of the working conditions of prostitutes. It is pointed out by Hughes (2004:3) that users of prostitutes do not distinguish between a prostitute who is a victim of human trafficking and a commercial sex worker. It appears that purchasers of sex workers prefer a person who is willing and able to render a sex service. In the meanwhile, a restricted and confined prostitute may appear more likely to render sex services with least resistance. Nevertheless, the clients interviewed in Coy, Horvath and Kelly, (2007:23) expressed a sense of discomfort with regards to accessing sex from caged women and appeared to sympathise with them rather than pursuing their motives. On the other hand, it is possible that a man who seeks sexual services for comfort would refuse sex from a caged person.

3.2.2 Profiteers of prostitution

Profiteers imply any person who generates profits from young women and girls forced into prostitution. It could be a club, brothel owners, pimps, massage parlours or owners of rental rooms, (SALRC 2009: 43-44). Brothel owners increase workforce by purchasing young women and girls from traffickers. In turn, once they are in their control, there are variety of ways in which brothel owners maximise profits from prostitutes (Gould & Fick 2007: 14) They determine the price for which the trafficked victim should charge from a trafficker (SALRC 2009: 45-46). Brothel owners may charge agency booking, weekly fees for advertising in newspapers and benefit from the misdemeanours committed by prostitutes whilst working within their confines such as coming to work late (Gould & Fick 2007:14). Young women and girls give payment received from the clients to owners of brothels. Pimps may purchase young women and girls for involuntary prostitution and could sell girls to other pimps to increase profits. Boyfriends and relatives could manage, by force young female relatives as well as intimate partners to work as prostitutes.

3.2.3 Publicity of prostitution

The publicity gained by commercial sex work from the print publications such as newspapers, academic journals and internet, may create an impression that selling and purchasing sex is an acceptable form of earning a living and accessing sexual services in South Africa and the world. Mansson, (2006: 90) is of the opinion that the mass production of sexualised images of prostitutes appearing in print media could be responsible for men's thinking that as long as one is willing to pay , sexual access is possible. To add on the glamour dimension are advertisements of girls selling sex appearing in newspapers. The research results revealed by Coy, et al. (2007:13) pertaining to the access routes from where men are likely to access prostitutes, illustrate that classified sections of newspapers appeared to be favoured by most users of prostitutes. To make the situation of prostitutes more attractive and humane are human rights organisations addressing challenges encountered by prostitutes whilst executing their duties. In the meanwhile, a legal and academic debate centred on abolishment or legalisation of prostitution is currently in progress in South Africa. Although the South African criminal justice system portrays ambivalent response towards sex work, until decriminalised, regulated or partially criminalised, according to section 20 of the Sexual Offences (Act 53 of 1957) it remains an illegal form of income generation. By responding to advertisements seeking girls for normal prostitution, girls could be lured into human trafficking for sexual exploitation.

Three variables can be drawn from the demand theory of prostitution: users, profiteers and publicity gained by prostitution from the public. All three variables point to the situational context of crime commitment. Prostitutes are portraying the victim; users and profiteers are depicting the offender; and publicity is signifying the opportunities available for prostitutes to exist in the public domain.

3.3 Victimological framework

There exist certain concepts within the Victimology paradigm explaining why certain women might be at greater risk of being victimised than others. Victim offender interaction, repeat victimization and lifestyle as a factor in crime victimization are key factors describing the nature of victimisation of human trafficking for involuntary prostitution. Relationships and interactions with traffickers, lifestyles and number of times women have been trafficked form an essential part of the process of human trafficking of women and girls for involuntary prostitution.

3.3.1 Victim criminal relationship

Victim proneness and victim-offender interaction are two factors explaining the vulnerability of young women and girls to human trafficking for involuntary prostitution through maintaining a relationship with a criminal. There are three categories of victims that could be prone to victimisation: innocent, precipitating and provocative victims. By interacting with criminals innocently through no fault of their own or by walking alone in the dark some individuals could be seen as precipitating their own victimisation. In addition, by exhibiting certain behaviours that could be seen as provocative by criminals, some women and girls are prone to victimisation, (Van den Hoven & Maree, 2005:61). The distance between the offender and the victim and the intentions of the offender and the nature of victim-offender interaction may increase the chances of victimization. It is pointed out by Van Den Hoven and Maree, (2005:61) that victims and the offender could have interacted closely before victimization occurred. Victim involvement in the events that led to victimisation could be identified. Either the victim had provoked or precipitated the victimization incident. However, it is a known fact that women and children are more likely to be victimised by a known person than by strangers. It remains to be seen whether young women and girls trafficked by close associates should be blamed or defended for the choices they have made. Victim proneness, precipitation and provocation will be used as the three factors which creates vulnerability of women and girls to human trafficking for involuntary prostitution

3.3.2 Repeat victimization

Repeat victimisation entails that victims of crimes are likely to be victimised either by different perpetrators or the same assailant during a limited time period, Van den Hoven & Maree, (2005:65). It is pointed out by Van den Hoven and Maree, (2005:66) that repeat victimization is likely to manifest into a cycle of violence. In turn, victims are likely to become abusers by replicating or modelling behaviour and actions perpetrated against them by human traffickers and recruit other women and girls to involuntary prostitution. According to Van den Hoven and Maree (2005:67) there are certain characteristics that make specific categories of people more prone to repeat victimisation than others: target vulnerability, target gratifiability and target antagonism. Target vulnerability is illustrated by victim's physical weakness, and psychological distress; whereas target gratifiability entails attributes that are attractive to the perpetrator. The presence of destructive impulses in one's personality could expose one to the risk of being victimised. With traumatic experiences unaddressed and future plans constrained by the sexual exploitation, and financial gains that could have been generated from prostitution, it is possible for former victims to either work independently as prostitutes or replicate the process by recruiting potential victims for their own benefit. Replication of human trafficking is conceptualised as second-wave trafficking. Repeat victimisation and second wave trafficking correlates. Victim vulnerability is created by repeat victimisation.

If the victim escapes human traffickers, the chances are that she might be re-victimised. The trauma experienced through the process of being victimised at first, may generate feelings of helplessness and could see the victim back with traffickers for involuntary prostitution. With trauma left unresolved, the dependence and bonding between human traffickers and victims still present in victims psyche and negative responses of the community and family members to the experience of the victim, the chances are that the victim could still be vulnerable to human traffickers. The human traffickers could be the only centre of acceptance that the victim could find comforting.

3.3.3 Lifestyle as a factor in crime victimization risk

Lifestyle risk model describes the risk of victimisations to be influenced by: personality of the potential victim, absence of a capable guardian, environment with which the potential victim resides and the daily routine activities which occupies the time of the potential victim, (Van den Hoven & Maree, 2005:63). According to Van den Hoven and Maree, (2005:63) the lifestyle activities which a potential victim participates determine the type of victimisation one could experience. For example, substance abuse may fuel the occurrence of interpersonal violence, whereas the use of the internet may expose children to inappropriate social networks, and participation in activities favoured by the deviant groups may pose danger to unsuspecting potential victims, (Van Den Hoven & Maree, 2005:65). Lifestyle activities such as substance abuse, internet use and equivalent groups are key factors explaining human trafficking of young women and girls for involuntary prostitution. It is possible for human traffickers to draw into the human trafficking ring known victims whose lifestyles intersect with theirs instead of selecting girls and young women from unfamiliar places. Such victims could least likely to report the perpetrators to the authorities thus rendering themselves easy prey to human traffickers, (Van den Hoven, 2005:65).

In summary, the victim's vulnerability to human trafficking for involuntary prostitution is created by victim's offender interaction, repeat victimisation and lifestyle pursued by the victim. In this regard, victim vulnerability is the variable selected to construct the integrated framework to explain human trafficking of women and girls for involuntary prostitution.

3.4 Consitutive theory

The central idea of constitutive criminology is that power and equality build socially constructed differences through which harm and deprivation is imposed on the subordinated group. The interconnectedness of societies which cannot be seen outside of cultural and structural contexts, determines the types of crimes that are likely to be perpetrated in specific geographical communities, (Lanier & Henry, 2004:321). Constitutive criminologists perceives criminals as excessive investors in crime who could use any means necessary to achieve the desired outcomes whereas a victim is often the disabled party who experiences pain, loss and denied humanity, (Lanier & Henry, 2004:323). As inter-continental trade agreements become a profitable way of conducting business, relations between countries expand. Business executives and non-governmental organizations travel frequently inter- and intra-continentally. With the world connecting on global scale immigration, traveling and tourism opportunities are now more accessible. Power and inequality, interconnectedness, investment in crime and loss and pain experienced by victims during the perpetration of human trafficking are identifiable variables important to build an integrated theoretical framework to explain human trafficking of young women and girls for involuntary prostitution. The constitutive cultural and structural contexts within which potential victims emanate are important to describe victim vulnerability to human traffickers.

3.5 Economic theory

Economic theory can be used to explain crimes, actions and behaviours which calculate the gains and benefits accrued from participating in a certain task. Economic theory of crime suggests that people make decisions to offend in ways that resemble their decisions made about other non-criminal activities, (Witt & Witte 2000:4, 6). The criminal might commit crime if the expected gains from legal work are less than the ones that are to arise from illegal work. The underlying principle of the economic theory is that, criminals commit crime because they have perceived the benefits from the crime to outweigh the possibility of being prosecuted and incurring costs, (Eagle & Betters, 2007:166; Persson & Siven, 2007:213). It is pointed out by Pratt (2008:44) as well as Witte & Witt (2002:2, 5) that individuals apply legal or illegal actions because of the expected utility from those acts and are influenced by the fact that the possibility of the expected gains from crime relative to earnings from legal work accentuates trafficking endeavours. Another component of this theory is that the lesser the punishment the more human trafficking progresses. The probability of being apprehended prosecuted and sentenced and the value of the expected punishment will determine the extent of the crime. That brings another economic dimension in human trafficking that is presented by McCray, (2006) in which he argues that certainty is more important than severity. Furthermore, McCray observes that the criminal in this regard would act like an economist and apply the image of a self-maximizing decision maker, carefully calculating his or her advantage, which might be different from an opportunist whose ill considered and reckless nature might get him into trouble. Certainty and severity of punishment are the two variables that are considered important for the formulation of an integrated theory.

In summary, once rational decisions have been made, vulnerability of potential victims ascertained and requirements of the demand considered the human trafficking process resumes. The human traffickers could by then have studied the legal response to human trafficking to ascertain the sanctions confronting them should they get caught by the criminal justice authorities. Hence, it is vital that a combination of public health, criminal justice, crime prevention and criminology conceptualized as Epidemiological Criminology should be considered to prevent, protect victims, and prosecute human traffickers of young women and girls for involuntary prostitution.

4. Theory integration

The formal structure for the integrated theory to explain human trafficking for involuntary prostitution takes the form of multiple causality. Human trafficking for involuntary prostitution is an outcome of a combination of multiple factors, (Lanier & Henry, 2010: 383). The concepts integrated for the construction of this framework are done at individual level of theory integration. There exists a mutual relationship between selected variables so that when combined one cannot function without the other. Decision making process will not occur without rational decision making which comprises of rational choice, the demand as well as victim vulnerability. Power and inequality, free will and lifestyle exposure are three concepts creating an opportunity for recruitment to occur. The interconnectedness of the world, victim precipitation as well as severity and certainty of punishment enable human traffickers to move the victims. Exploitation of victims is made easier by the ambivalent attitudes expressed by society towards the victims as well as the costs and benefits generated from prostitution. The investment in crime and profiteers from the involuntary prostitution give effect to the harbouring and transfer of victims. If victims exit the human trafficking process alive, the loss and pain endured from the experience might likely to influence them to either return as recruiters or work independently as prostitutes. To explain the factors contributing to human trafficking of women and girls for sexual exploitation: the variables below have been selected from the theories described above. Each set of variables correlate with each stage of human trafficking from recruitment to loss and pain.

4.1 Variable 1: Decision making

There are three causal or explanatory factors to the decision making process for committing the crime of human trafficking of young women and girls for involuntary prostitution: **Free will, the demand and victim vulnerability.** Decision making process becomes an independent variable because it precedes the free will, demand and victim vulnerability. Dependent variables are drawn from rational choice, victimology and demand theories to explain the decision making process of human traffickers prior the commitment of human trafficking of young women and girls for involuntary prostitution. The assumption in this regard is that human traffickers will not observe the demand, victim vulnerability or free will in the absence of a decision to commit the crime. The demand will not express interest if the promise of their needs to be fulfilled is not relayed by human traffickers. Young women and girls will not be lured into prostitution if traffickers do not rationally create falsehoods that could see one accepting a dubious job in a different location. Essentially, there exists a mutual relationship between free will, the demand as well as victim's vulnerability.

Certainly, three activities happen before human trafficking for involuntary prostitution occurs. Firstly, human traffickers choose this crime rationally by calculating the costs and benefits to be generated from selling young women and girls as prostitutes to pimps or brothels. Furthermore, the ambivalent social attitudes and approaches towards prostitution clear the path for human traffickers to commit the crime. Secondly, they could establish the financial capabilities of the demand and the type of girls preferred by the potential users of victims. Lastly, by ascertaining vulnerability, they determine the easiness with which they could assess the specific type of girls preferred by the demand. For example, within an organised crime context, human traffickers could decide on human trafficking once they have ascertained the leeway paving the way for the commitment of drug dealing such as a drug courier, officials to corrupt and the demand for the drugs.

On the contrary, Joubert (2008:112), Lanier & Henry (2010:81) as well as Brown et al. (2007:219) cautions against contending that rational decision making might be preceded by other variables prior the commitment of crime for the reason that some crimes could be committed impulsively. However, human trafficking for involuntary prostitution involves not only the victim, opportunity and the perpetrator. It involves different locations, spaces, participants and routes; namely victim's country of origin, country of transit or destination from where victimisation might occur and potential users of victims. Nevertheless, rational choice theories are not enough to explain the decision making process involved prior the commitment of human trafficking for involuntary prostitution. The crime involves the situational context which puts victims in the position of vulnerability, better explained by the victimological theories and the demand theories that provide insights into the factors behind the use of prostitutes, ambivalent attitudes towards prostitution as well as the purchasing of young girls and women by brothel owners. Once the human traffickers have decided on the crime, identified victims, analysed the situational factors that could lure victims by measuring the distance between themselves and potential victims, recruitment phase resumes.

4.2 Variable 2: Recruitment

A casual relationship exists between **power and inequality, free will and lifestyle exposure** creating opportunities for the recruitment process to occur. Recruitment becomes an independent variable and power and inequality, free will and lifestyle exposure are dependent variables. The dependent variables to describe the manner in which human traffickers recruit victims for involuntary prostitution are derived from the constitutive criminology, rational choice and lifestyle exposure model. For the recruitment phase to occur, there ought to exist unequal power differences between the trafficker and the victim. The victim should need to access opportunities for economic advancement. The traffickers may come up with a strategy that could help advance the victim. Social position, psychological well-being, self-realization and actualisation, (Lanier & Henry 2010) of victims should condition them to accept the offers made by human traffickers. In addition the victim should pursue a lifestyle that positions her closer to the traffickers. Essentially power and inequality influences the free will to recruit potential victims based on the lifestyle personality of the victim. With little or no knowledge of victims' lifestyle, gender, socio-economic status, ethnicity and race per preference of the demand, the traffickers will not recruit victims.

4.3 Variable 3: Transportation

Constitutive criminology, victimology and economic crime explain the easiness in which human traffickers move victims from one location to the other. The process of transporting victims from one location to the other occurs because of the global interconnectedness, victim precipitation as well as certainty and severity of punishment. Human trafficking for involuntary prostitution is an outcome of interactions between human beings across the world, made possible by the interconnectedness of human beings, social and organisational structures, (Lanier & Henry, 2010). The mobility of people from one country to the other, accessibility of the victim to the traffickers as well as the ineffectiveness of the criminal justice system in arresting, prosecuting and convicting human traffickers, could explain the transportation stage of the trafficking process. The same modes of transportation used to transport non-victims of human trafficking are used by human traffickers to move victims.

A correlation exists between transportation and interconnectedness. These two variables occur simultaneously. Interconnectedness may appear to take the place of an independent variable. If the world were a closed entity exempt from intercontinental trade agreements, foreign investment and technological advancements, human traffickers would not transport victims in abundance in the manner that they are currently moving. However, the primary intention to be achieved in this regard is the transportation of human trafficking victims. Transportation takes precedence over interconnectedness. The global connection just creates ease with which human traffickers conduct their business.

Victim precipitation as well as severity and certainty of punishment become third variables which strengthens the relationship between interconnectedness and transportation. The acceptance of employment opportunities as well as travelling to another location may seem to create an impression that the victim agrees with the plans of the human trafficker. However, the intended outcome should not benefit the victim but the traffickers. The weak response of the criminal justice system to the crime reinforces the transportation of victims.

4.4 Variable 4: Exploitation

Exploitation of victims of human trafficking for involuntary prostitution would occur despite the publicity created by the demand, victims' proneness and costs and benefits. It is pointed out by Fick, (2005); Gould and Fick, (2007) as well as SWEAT, (n.d) that exploitation of prostitutes is prevalent: clients, police officers and the public prone to victimise prostitutes. Exploitation becomes an independent variable which occurs alongside the bad and good publicity received by the industry, victim proneness to violence as well as the monies that are generated by the pimps and traffickers from prostitution. The assumption is that victims of human trafficking for involuntary prostitution are exploited because: Firstly, by being in the industry they are prone to victimisation. Secondly human traffickers intend to generate huge returns from the business of prostitution. Lastly, the space occupied by victims makes them least likely to avoid exploitation.

The demand theory, victimology and rational choice theory concepts explain the manner in which human traffickers, users and profiteers of prostitution exploit sex workers. The publicity that media advertisements, academic debates and representations by human rights organisations give to the sex work industry creates an impression that prostitutes are bodies to be exploited. A client, who purchase sex services for the purpose of displacing anger, may have psychologically and biologically perceived prostitutes as provocative bodies expecting to be violated. The manner in which the prostitutes dress, the precipitative words they use to solicit clients and the media images glamorizing the sex business could provoke a violent client to use violence towards sex workers. Whilst the cost and benefit analysis calculated by human traffickers prior commitment of crime allows exploitation of trafficked victims to occur. Human traffickers may not risk getting caught, or share financial gains or the skills needed to successfully commit the crime with the victims for the reason that they will not obtain the goals set prior crime perpetration.

4.5 Variable 5: Harbouring and transfer

Harbouring and transfer are vital activities causing brothel owners, clients as well pimps to profit and invest in the human trafficking of young women and girls for involuntary prostitution. The extent to which they are willing to keep victims and use them to generate income is one of the factors driving the human trafficking process. However, they vary in the manner in which they retain victims for victimisation. Harbouring and transfer are the concepts derived from demand and constitutive theories to comprehend the motives behind the confinement and selling of victims by human traffickers. The victims are sold because the crime in which the traffickers have invested generates profits. In essence, harbouring and transfer become an independent variable: taking precedents over profits and investment in crime. Harbouring and transfer causes profiteers and investment in crime. Profiteers and investment in crime become dependent variables. Brothel owners buy victims because they are available by means of confinement by traffickers. In turn, participating in crime for prolonged time becomes an investment in crime.

The longer the traffickers keep victims confined and performing involuntary sex work, the more profits they generate from the crime. The investment in the human trafficking of young women and girls for involuntary prostitution persuade traffickers to apply every means necessary to keep control of the victims. Profits are a necessary prerequisite for the harbouring and transfer of the victims. The greater the profits traffickers generate, the more creative traffickers' become in the methods they use to invest in the crime.

4.6 Variable 6: Loss and pain

The economic, physical, sexual and psychological losses and pains encountered during the process of human trafficking for involuntary prostitution could create second wave trafficking. Derived from constitutive criminology the concept of loss and pain resonate with the occurrence of second wave trafficking. Constitutive criminology postulates that crime is "power to deny others their ability to make a difference", (Lanier & Henry, 2010). According to Lanier and Henry (2010) victimisation is a sign of disrespect to the victims – when victimised victims are prevented from interacting with others, transform themselves and better their economic situation. Victims do not only lose income and integrity but their sense of self-worth and value. By being forced involuntarily to perform sex work, victims could be severely traumatised and physically damaged. In turn, if the trauma of loss and pain is left unresolved they might likely recruit other girls into human trafficking for involuntary prostitution. Thus, the cycle of violence and victimisation is displaced upon a second group of victims.

5. Theory synthesis

An integrated framework to explain and describe the process of human trafficking was described in this paper. To construct an integrated framework this paper coalesced variables from different disciplines to explain the occurrence of human trafficking. To synthesize, human traffickers and the demand, each play an instrumental role in the victimization of trafficked women. Identified are unguarded victims - seen as attractive targets who are innocent or facilitating the process, cooperative even and whose countries may be experiencing some structural constraints - to supply the needs of the demand. Traffickers follow an economic approach to perform human trafficking by calculating the strength of punishment against the financial incentives likely to be drawn from the process of human trafficking. By taking advantage of globalization issues, they commit the crime. Human trafficking is a constitutive crime that involves role players from diverse populations and professional backgrounds. As the world becomes more interconnected human trafficking is increasing rapidly. In addition, constitutive criminology points out that, victims are always at the receiving end of the human trafficking process. They experience, emotional loss, suffering and dislocation during and after the trafficking process.

There are currently many varied approaches applied by academics and researchers' to explain and respond successfully to the challenges faced by victims and potential victims of human trafficking for involuntary prostitution. The expanding area of Epidemiological Criminology or "EpiCrim" might be useful to devise strategies to respond effectively to human trafficking of young women and girls for involuntary prostitution (Lanier, Pack & Akers 2009). EpiCrim emphasizes the need to provide public health, justice, victim support and investigation to victims of human trafficking (Akers & Lanier, 2009). Public awareness on human trafficking for potential victims is a focus of concern for EpiCrim adherents. The purpose of this approach is to ensure that the cycle of violence is not repeated. By providing justice and services to victims, redress for the crime suffered is conducted.

5.1 Policy suggestions: The South African experience

There are few chances of getting caught and convicted for the crime of human trafficking in South Africa (SA). Human Trafficking is not even a legally defined crime in SA as the country is still in the final stages of approving the Human Trafficking Bill. In the meantime, some sections of legislations such as Criminal Law (Sexual Offences and Related Matters) Amendment Act (32 of 2007), Children's Act (38 of 2005), Prevention of Organized Crime Act, (121 of 1998) and Immigration Act, (13 of 2002) are used to respond to the crime. Human trafficking accused have been appearing in court to defend allegations of sexual offences, child maltreatment and participation in organized activities, with most being sentenced under the Prevention of Organized Crime Act (121 of 1998). Dryden et al. (2001:6) are of the opinion that the inevitability of crime commission might be obvious in a situation where the legal sanctions are not available to respond to crime or if the sentencing procedures is not consistent. As a consequence of the lack of an approved legislation to respond and combat human trafficking for all purposes, South Africa has been put on tier two by the United Nations Convention against Transnational Crime: meaning that whilst contingency measures are underway to respond promptly to this crime, the process is slow. Non-governmental organizations and some research institutions perform prevention and protection duties especially for victims and potential victims of human trafficking.

5.2 Consequences of human trafficking for victims

Negative consequences outweigh positive consequences of human trafficking for trafficked victims. Human rights violations and deterioration of mental and physical health are impairments experienced by victims of human trafficking during and after the process of trafficking. Free will, human dignity and an ability to make decisions are some of the human rights violations experienced by victims at the hands of human traffickers (Hughes, 2004; Kinnu, 2006:24; Phinney, 2006:4; Timoshkina & MacDonald, n.d:17). Trafficked young women and girls often have traveling documents confiscated by human traffickers, which render them incapable of seeking assistance from the authorities at the host country (Jordan, 2002:35; Simic, 2004:25; Timoshkina & MacDonald, nd: 17). They are often deprived of basic medical and mental health care; shelter that is not a form of prison or detention; protection from traffickers; access to information on legal rights and to attorneys or advocates; financial or other assistance, for example food, clothing and telephone calls as well as a means to return home safely. Having proof of identity, traveling documents or a work permit, would provide the victim permission, short stay or employment, at the host county.

The often confining process of human trafficking does not leave victims unscathed; trafficked victims could experience nightmares, depression, anxiety, sexually transmitted infections, back aches and other mental health related ailments. Epi-criminological framework is recommended to respond to victim experiences, rehabilitate human traffickers and assist other persons that could indirectly be affected by human trafficking such as families and relatives of victims.

6. Summary

The purpose of this paper was to construct an integrated framework for explaining and describing human trafficking of women and girls for involuntary prostitution. Variables from existing theories were drawn to formulate a comprehensive view of the process followed during the commitment of this crime. The basic argument is that, since human trafficking involves a high contingent of role players from a variety of backgrounds, a single explanation for its cause cannot exist. Essentially, human trafficking responses should consider the importance of an integrated framework in order to effectively, prevent, prosecute suspects and protect victims of human trafficking of women and girls for involuntary prostitution.

7. References

Akers. T., & Lanier, M. (2009). Epidemiological Criminology: Coming Full Circle. *American Journal of Public Health*, 99 (2), 1-6.

Bales, K. (2009) Presentation at the University of Central Florida, January 15, 2009.

Bruckett, C., & Parent, C. (2002). *Trafficking in human beings and organized crime: A literature Review*. Ottawa, ON, Canada: RCMP. [O] Avaialble:http://www.rcmp-grc.gc.ca Assessed 14/03/2008.

Brown, S.E., Esbensen, F., & Geis, G. (2007). *Criminology: Explaining crime and its context*. Sixth Edition. Cincinnati: Anderson Publishing.

Children's Act 38 of 2005. *Government Gazette*, (28944). Pretoria: Government Printer.

Coy, M., Horvath, M., & Kelly, L. (2007).*'Its just like going to the supermarket': Men buying sex in East London. Report for safe exit*. London: London Metropolitan University.

Criminal (Sexual Offences and Related Matters) Amendment Act 32 of 2007. *Government Gazette*. Pretoria: Government Printer.

Davis, L. (2005). Theoretical approaches and perspectives in Victimology. In Davis, L. & Snyman, R. *Victimology in South Africa*. Pretoria: Van Schaik Publishers (pp 35-52).

Dryden Witte, A., & Witt, R. (2001). What we spend and what we get: Public and private provision of crime prevention and criminal justice. *Fiscal Studies*, 22 (1), 1-40.

Eagle, J.G., & Betters, D.R. (1998). The endangered species act and economic values: A comparison of fines and contingent valuation studies. *Ecological Economics*, 26 (1998), 165-171.

Fick, N. (2005). *Coping with stigma, discrimination and violence: Sex workers talk about their experiences.* Cape Town: SWEAT.

Groom, T.M., & Nandwani, R. (2006). Characteristics of men who pay for sex: A UK sexual health clinic survey. *Sexually Transmitted Infections*, 82, 364-367.

Gould, C., & Fick, N. (2007). *Report to the South African Law Commission: Preliminary research findings or relevance to combat trafficking in persons and legislation pertaining adult prostitution.* Cape Town: SWEAT.

Henry, S., & Milovanovic, D. (1998). *Constitutive Criminology: Beyond Postmodernism*. London: Sage Publications.

Hughes, M. D. (2004). *Best practices to address the demand side of sex trafficking*. Rhode Island: University of Rhode Island.

Immigration Act 13 of 2002. *Government Gazette*, (26901). Pretoria: Government Printer.

Jordan, A. (2002). Human rights or wrongs? The struggle for a human rights-based response to trafficking in human beings. In Masika, R. (Ed). *Gender, Trafficking and Slavery.* Oxford: Oxfam Publishers.

Joubert, S. (2008). Contemporary theoretical explanations for youth misbehaviour. In Bezuidenhout, C. & Joubert, S. *Child and Youth Misbehaviour in South Africa: A holistic approach.* (pp 108-122) Pretoria: Van Schaik Publishers.

Kinnu, G. (2006). *From Bondage to Freedom: An analysis of International Legal Regime on Human Trafficking*. [O] Available:http://www.nhrc.nic.in

Lanier, M., & Henry, S. (2004). *Essential Criminology.* 2nd edition. United States of America: Westview Publishers.

Lanier, M., & Henry, S. (2010). *Essential Criminology.* 3rd edition. United States of America: Westview Publishers.

Lanier, M., Pack. R. P. & Akers. T. (2009). Epidemiological Criminology: Drug use among African American gang members. *Journal of Correctional Health Care*, x:xx-xx.

Macleod, J., Farley, M., Anderson, L., & & Golding, J. (2008). *Challenging men's demand for prostitution in Scotland: A research report based on interviews with 110 men who bought women in prostitution*. Glasgow: Women's Support Project.

Mansson, S.A. (2006). Men's demand for prostitutes. *Sexologies*, 15 (2006): 87-92.

McCray, J. (2006). *Dynamics and the economic theory of crime. Criminology and Economics Summer Workshop.* University of Michigan: June 5.

Persson, M. Siven, C-H. (2007). The Becker Paradox and type I versus type II Errors in the Economics of crime. *International Economic Review*, 48 (1), 211-233.

Phinney, A. (2006). *Trafficking of women and children for sexual exploitation in the America's*. [O] Available:http://www.planetwire.org/wrap/files.fcgi/2369_trafficking_paper.htm Assessed 25/03/2008.

Pratt, T.C. (2008). Rational Choice Theory, Crime Control Policy, and Criminological Relevance. *Journal unknown*, 7(1), 43-52.

Prevention of Organized Crime Act 121 of 1998. *Government Gazette*, (19553). Pretoria: Government Printer.

Simic, O. (2004).*Victims of trafficking for forced prostitution: Protection mechanisms and the right to remain in the destination countries*. Switzerland: Global Commission on International Migration.

Snyman, R. (2005). Overview of and concepts in Victimology. In Davis, L & Snyman, R. (Eds.) *Victimology in South Africa.* Pretoria: Van Schaik Publishers.

South African Law Reform Commission (SALRC). (2009). Discussion Paper 0001/2009, Project 107: Sexual Offences Adult Prostitution. Pretoria: SALRC.

Stop Violence Against Women. (SVAW). (2008). *Factors that contribute to trafficking in women*. [O] Avaialable:
http://www.umn.edu/humanrts/svaw/trafficking/explore/3factors.htm
Assessed 25/03/2008.

SWEAT. [Sa]. *Policing sex workers: A violation of human rights*. Cape Town: SWEAT.

Timoshkina, N., & MacDonald, L. [Sa]. Defining Trafficking in Women. [S1:sn).

Truang, T.D. (2006). *Poverty, gender and human trafficking in Sub-Saharan Africa: Rethinking best practices in migration management* [O] Available:http://www.popline.org/docs Assessed 14/03/2008.

Truong, T. (2001). *Human Trafficking and Organized Crime. Institute of Social Studies Working Papers Series no.339.* [O] Avaialable:http://www.iss.nl. Assessed 25/03/2008

United Nations. (2000). *Protocol to prevent, suppress and punish trafficking in persons, especially women and children, supplementing the United Nations convention against transnational organized crime.* [O] Available:http://www.un.org.za Assessed 26/02/2008.

Van den Hoven, A., & Maree, A. (2005). Victimization risk factors, repeat victimization and victim profiling. In Davis, L. & Snyman, R. *Victimology in South Africa* (pp55-71). Pretoria: Van Schaik Publishers.

Van Impe, K. (2000). 'People for sale': The need for a multi-disciplinary approach towards human trafficking. *International Migration, Special Issue*, 2000 (1):113-131.

Witte, A.D., & Witt, R. (2002). Crime Causation: Economic Theories. In *Encyclopaedia of Crime and Justice*, 1 (302-308) New York: Macmillan.

Edited by Jay Maddock

Human behavior accounts for the majority of morbidity and premature mortality throughout the world. This book explores several areas of human behavior including physical activity, nutrition and food, addictive substances, gun violence, sexual transmitted diseases and more. Several cutting edge methods are also examined including empowering nurses, community based participatory research and nature therapy. Less well known public health topics including human trafficking, tuberculosis control in prisons and public health issues in the deaf community are also covered. The authors come from around the world to describe issues that are both of local and worldwide importance to protect and preserve the health of populations. This book demonstrates the scope and some of the solutions to addressing today's most pressing public health issues.

Photo by agsandrew / iStock

IntechOpen

ISBN 978-953-51-4295-9