IntechOpen

eHealth

Making Health Care Smarter

Edited by Thomas F. Heston

EHEALTH - MAKING HEALTH CARE SMARTER

Edited by **Thomas F. Heston**

eHealth - Making Health Care Smarter
http://dx.doi.org/10.5772/intechopen.71820
Edited by Thomas F. Heston

Contributors

Feras M Ghazawi, Ivan V. Litvinov, Steven Glassman, Denis Sasseville, Sandeep Reddy, Geletaw Sahle Tegenaw, Loc Nguyen, Thu-Hang Ho, Truong-Duyet Phan, Catherine Demers, Sahr Wali, Karim Keshavejee, Sinchai Kamolphiwong, Thossaporn Kamolphiwong, Verapol Chandeeying, Soontorn Saech, Alexandre Serra Barreto, Dominik Alscher, Daniel Luna, Laura Gambarte, Carlos Otero, Julia Frangella, Thomas F. Heston

Notice

First published in London, United Kingdom, 2018 by IntechOpen
eBook (PDF) Published by IntechOpen, 2019
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, The Shard, 25th floor, 32 London Bridge Street
London, SE19SG – United Kingdom
Printed in Croatia

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

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

eHealth - Making Health Care Smarter
Edited by Thomas F. Heston

p. cm.

Print ISBN 978-1-78923-522-7

Online ISBN 978-1-78923-523-4

eBook (PDF) ISBN 978-1-83881-621-6

Meet the editor

Dr. Thomas F. Heston is a practicing physician and clinical associate professor at the Elson S. Floyd College of Medicine in Spokane, Washington, USA. As a nuclear medicine resident in the early 1990s, he researched the application of neural networks in transplant medicine. He subsequently started the first nuclear medicine teleradiology network to rural northern Idaho and in the process single handedly developed the picture archiving and communication system for the network. He then went on to study positron emission tomography at Johns Hopkins University. He became an assistant professor (adjunct) in their Department of Radiology and a medical director for Johns Hopkins International. His current research focus is on the application of blockchain technology to health care.

Contents

Preface

This first edition of *eHealth: Making Health Care Smarter* provides readers with an introduction to the wide-ranging, important applications of mathematical algorithms and electronic technology to the improvement of health and the prevention of disease. The field of eHealth is so large that any textbook, no matter how comprehensive, will exclude important areas of this revolution in medicine. The chapters contained within this book provide the reader with insights into the focus of leading world scientists in the field of eHealth. As a practicing physician for over 25 years, I have witnessed medicine as it existed before the widespread growth of the Internet and smartphones. The improvements brought forward by eHealth are truly revolutionary.

Thomas F. Heston, MD, FAAFP, FASNC, FACNM
Clinical Associate Professor
Elson S. Floyd College of Medicine
Washington State University
Spokane, Washington, USA

Introduction

Chapter 1

Introductory Chapter: Making Health Care Smart

Thomas F. Heston

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.78993

1. Introduction

The age of eHealth and Smart Medicine is upon us, but what exactly does this mean? As technology advances, we are able to create electronic devices that collect and analyze data, electronic communication methods that alert health care providers immediately when adverse events arise, and electronic algorithms that help automate and speed up clinical decision-making.

A primary leader in smart medicine is the use of wearable technology. These electronic devices enable the collection of important medical data. Combining wearable devices such as heart rate monitors, pulse oximeters, and sleep monitors with blockchain technology allows this important patient information to be recorded accurately, remain immutable over time, and interact with algorithms designed to improve medical diagnosis and treatment. Wearable technology is already well developed. Making this technology interoperable with electronic medical records in a manner allowing smart execution of health care protocols becomes possible with the use of blockchain technology.

Satoshi Nakamoto set forth the initial implementation of blockchain technology in the white paper "Bitcoin: a peer-to-peer electronic cash system" in 2008 [1]. This white paper presented a method to create an Internet-based currency that did not require a trusted third-party intermediary such as a bank, government, or Federal Reserve. Instead of using a third-party intermediary, the blockchain method utilized computers hooked up to the Internet to confirm transactions in a manner that would prevent malicious hacking, cheating, or double-spending. Bitcoin was subsequently created, with the first transaction occurring in January, 2009. Nakamoto's blockchain method serving as the foundation for bitcoin has proven to be widely successful, with the market capitalization of bitcoin as of early 2018 equal to approximately $150 billion USD.

The backbone for bitcoin is a simple blockchain of transactions that is immutable and secure due to a global distributed network of computer nodes (also known as miners) that confirms new transactions and secures old transactions. This distributed ledger technology works well, powering about $2 billion USD in transactions per day, with a total number of financial transactions to date of over 300 million [2, 3]. The success of bitcoin has created a wide expansion of blockchain technology, to the point where distributed computers around the world now confirm smart contracts [4], provide cloud storage [5], and facilitate communications between small devices (e.g., wearable wrist health bands) that make up the Internet of Things [6].

Through the integration of electronic devices with blockchain technology, the utility of wearable monitors increases tremendously [7]. By creating an immutable, trusted ledger of patient data, blockchain technology not only allows monitors to trigger human responses but also collects important physiologic information that can be analyzed later by both human doctors and also by "digital doctors," i.e., smart algorithms that would trigger actions based upon the input. In the blockchain world, these smart algorithms that trigger actions are called smart contracts [8].

Digital doctors can serve multiple purposes. First of all, alarms set off by existing monitors in most hospitals can be missed for example when the medical ward is busy making hearing the alarm more difficult. Monitors displaying bells or popups are only effective when a human is actively monitoring the screen in a focused, non-distracted way. Digital doctors, however, act according to algorithms, which execute instantaneously. Digital doctors do not get distracted, they do not require sleep, and they have an infinite attention span.

What can these digital doctors do, and why do they require blockchain technology? First of all, digital doctors can instantly initiate codes. For example, a "code sepsis" can instantly be initiated whenever a patient's vital signs become unstable; a "rapid response code" could be instantly initiated whenever the cardiac monitor displayed a malignant arrhythmia. In some cases, these digital doctors could act spontaneously without human intervention (e.g., this is done with wearable insulin pumps and implanted cardiac defibrillators), and in other cases they could trigger initiation of a medical treatment protocol that would require physician review before implemented.

The key to digital doctors becoming useful and effective is trustworthy, accurate, immutable, and private data. Medical care requires accurate collection of patient health data. Scans must be done properly, blood tests must be processed appropriately, and real-time monitors must be calibrated. This is where blockchain technology can really help, because it allows the collection of data in a prompt manner that can be trusted and immutable. Recording data for digital doctors in a centralized database would result in a system that was vulnerable to a single point attack, whether it be an electricity failure or human hacker. Blockchain technology, on the other hand, would make the data more interoperable by ensuring it is readily accessible to digital doctors. It would make the data more reliable through blockchain consensus mechanisms that would be strongly resistant against hacking. It would also make the data easier to audit for quality improvement purposes. Finally, using cryptography inherent in blockchain technology, patient confidentiality is prioritized [9].

Blockchain technology creates trustworthy data that is reliably stored, easily accessed, and resistant to corruption. Wearable technology such as heart rate monitors, bed monitors, and pulse oximeters collect important information that when entered into a blockchain ledger can be processed by digital doctors that not only can be programmed by expert physicians, but can ultimately learn and improve through artificial intelligence. As we have seen in the home, the Internet of Things (IOT) has led to considerable advances in the creation of smart homes. Now, this technology is being applied to monitoring health with wrist monitors, blood glucose monitors, temperature monitors, and more. The time is right for not only having smart homes, but having smart hospitals. IOT along with blockchain technology is leading the way.

Author details

Thomas F. Heston

Address all correspondence to: tom.heston@wsu.edu

Elson S. Floyd College of Medicine, Washington State University, Spokane, Washington, USA

References

[1] Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System [Online]. 2008. Available from: https://bitcoin.org/bitcoin.pdf. [Accessed: 7 June 2018]

[2] Bitcoin Now Processes $2 Billion Worth of Transactions per Day, A 10x Increase in 2017 [Internet]. Available from: https://www.forbes.com/sites/ktorpey/2017/11/20/bitcoin-now-processes-2-billion-worth-of-transactions-per-day-a-10x-increase-in-2017/#6cd1eed32fba [Accessed: May 21, 2018]

[3] Total Number of Transactions [Internet]. Blockchain.info. Available from: https://blockchain.info/charts/n-transactions-total?timespan=all [Accessed: May 20, 2018]

[4] Buterin V. Visions, Part 1: The Value of Blockchain Technology [Internet]. Ethereum Blog. 2015. Available from: https://blog.ethereum.org/2015/04/13/visions-part-1-the-value-of-blockchain-technology/ [Accessed: July 13, 2017]

[5] Harper C. What is Siacoin? A Beginner's Guide to Decentralized Cloud Storage [Internet]. CoinCentral. 2018. Available from: https://coincentral.com/siacoin-beginner-guide/ [Accessed: May 21, 2018]

[6] Mark. What Is IOTA? [Internet]. The Merkle. 2018. Available from: https://themerkle.com/what-is-iota-cryptocurrency/ [Accessed: May 19, 2018]

[7] CREDITS. Blockchain is the key to the Internet of Things in medicine [Internet]. 2018. Available from: https://medium.com/@credits/blockchain-is-the-key-to-the-internet-of-things-in-medicine-71668652d88b [Accessed: May 09, 2018]

[8] Shaik K. Why blockchain and IoT are best friends [Internet]. Blockchain Unleashed: IBM Blockchain Blog. 2018. Available from: https://www.ibm.com/blogs/blockchain/2018/01/why-blockchain-and-iot-are-best-friends/ [Accessed: May 08, 2018]

[9] Snell E. Why Blockchain Technology Matters for Healthcare Security [Internet]. 2016. Available from: https://healthitsecurity.com/features/why-blockchain-technology-matters-for-healthcare-security [Accessed: April 12, 2018]

Fundamentals

Chapter 2

Terminology Services: Standard Terminologies to Control Medical Vocabulary. "Words are Not What they Say but What they Mean"

Daniel Luna, Carlos Otero, María L. Gambarte and Julia Frangella

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75781

Abstract

Data entry is an obstacle for the usability of electronic health records (EHR) applications and the acceptance of physicians, who prefer to document using "free text". Natural language is huge and very rich in details but at the same time is ambiguous; it has great dependence on context and uses jargon and acronyms. Healthcare Information Systems should capture clinical data in a structured and preferably coded format. This is crucial for data exchange between health information systems, epidemiological analysis, quality and research, clinical decision support systems, administrative functions, etc. In order to address this point, numerous terminological systems for the systematic recording of clinical data have been developed. These systems interrelate concepts of a particular domain and provide reference to related terms and possible definitions and codes. The purpose of terminology services consists of representing facts that happen in the real world through database management. This process is named Semantic Interoperability. It implies that different systems understand the information they are processing through the use of codes of clinical terminologies. Standard terminologies allow controlling medical vocabulary. But how do we do this? What do we need? Terminology services are a fundamental piece for health data management in health environment.

Keywords: terminology server, interface vocabulary, controlled vocabularies, semantic interoperability, standard terminology

1. Introduction

Recently, major healthcare stakeholders around the world have emphasized on the importance of establishing electronic health records (EHR) for all health care institutions. Their goals for doing so include increasing patient safety, reducing medical errors, improving efficiency and reducing costs [1, 2]. Everyday practical data entry, presentation and document retrieval for clinical tasks must be taken into account, so that the differences between the needs of users and the needs of available software's are addressed. Data entry is an obstacle for the adoption of EHR with structured data method and the acceptance of healthcare providers, who prefer to document healthcare findings, processes and outcomes using unfettered "free text" or narrative text in natural language [3]. Natural language is huge and very rich in details but at the same time ambiguous, having great dependence on context, it uses jargon and acronyms and it lacks of rigorous definitions.

1.1. The importance of narrative

Free text narrative formats allow physicians to share complex ideas in an efficient and effortless manner. It use in electronic health records allows them to synthesize facts and to point a full picture rich with meaning that it might be easily interpreted by other health care providers [3]. Between physicians` register motivation, the main one is their own use of the information. Many current systems that provide EHRs use template-based system in order to capture structured data elements in databases. Structure data entry does not support the expressiveness and flexibility to which clinicians are accustomed, and it can be difficult to interpret and reconstruct meaning from structure data due to loss of contextual information [4]. To represent medical knowledge, it is necessary to represent patient's data from different sources including: problem list and sometimes progress notes, procedures, medication list, labs and complementary tests results, social determinants of health environmental information, people's decisions about health and medical treatments, genomics and proteomics, etc. As a result, ambiguities must be resolved and vocabulary standardized.

1.2. The need of a standard codification system

To accomplish these, EHR should capture the clinical data in a structured and preferably coded format. Looking at the definition of codifying, we found "To reduce to a code" [5]. Codes are usually numeric or alphanumeric. In order to represent facts that happen in the real world to be managed in a database, the need of a standard codification system (SCS) arise. Evans et al. stated that the medical community required a "common, uniform, and comprehensive approach to the representation of medical information" [6].

This SCS should be able to capture clinical findings, index medical records, index medical literature and represent medical knowledge, etc. Provided that possible, the codification should be one-to-one: one term should only exist for a given object. Each term should describe only one object. The aim is to avoid ambiguity through polysemy or homonymy [5].

In fact, many SCS have been proposed but their adoption has been slow and incomplete. System developers generally indicate that, while they would like to make use of standards, they cannot

find one that meets all their needs. Each author who expresses a need for a controlled vocabulary does so with a particular purpose in mind, so there are also multiple characteristics that it should accomplish [7, 8]. Because of all the reasons mentioned before, for a long time, there has been a discussion regarding the use of free text versus structured text for data entry in EHR that later must be codified. Free text has the advantage of allowing health care providers to express themselves freely, but as disadvantage it has the need for an arduous codification process to allow further analysis. Structured text allows a quick codification process but has the disadvantage of being time consuming for the physician and contains expressions to the level of detail of the selected entry terminology [9]. It has been suggested that tension between clinical usability and meticulous knowledge representation may result from a fundamental conflict between the needs of humans and those of computer programs that use terminologies [3].

1.3. Primary coding versus secondary coding

Ideally, clinical data should be coded by the practitioner at the time of the consultation, which is known as primary coding, so they can utilize their knowledge of the patient situation while being aware of the limitations set down by the selected classification system [10]. However there are several practical difficulties in setting primary coding. It is time consuming for the practitioner and requires major efforts in their training, to ensure that the same code would be chosen in the same situation by different physicians [10]. It also limits the physician's expression in the registry and creates high levels of resistance in its use. Enforcing mandatory as opposed to optional modifier codes results in lower rates of incomplete coding [11]. One answer to this problem is centralized secondary coding, where a reduced number of trained persons codify the narrative text recorded by the physicians taking care of the patients. Centralized secondary coding by non medical coders had proved to be reliable and can be used for coding medical problems from an electronic problem-oriented medical record [10]. As regards as the coding tool, manual coding versus computerized coding, it has been demonstrated that the use of a computerized coding tool can save time and result in higher quality coding. A study that compares both of them had shown that manual coding takes 100% longer [11]. It is fundamental to contemplate that time spent on coding may be underestimated when we look at individual coding times instead of looking at the whole task of processing a clinical scenario [11]. The completeness of coding had also been demonstrated that can be improved using a computerized coding tool [11]. A step-forward option is to achieve text-autocoding, allowing free or narrative text entry together with dynamic interaction of the information system at the time of entering data.

As a result, the challenge consists on finding the complex balance between the freedom of use of free text and the benefits of structured text for data entry in EHR. In order to answer to this need, interface terminology and a terminology server were developed. It is crucial to highlight that communication is successful only if the sender and the receiver know both, the language (code) and the context. Notice the importance of the context, which must be read in an identical manner for both parties involved [5].

1.4. But again, why do we need to codify in electronic health records?

Some of the aims to coding in EHR are:

- to support health services research: this system can promote quality of care by providing a link to medical knowledge and current publications that can be used for outcome measurement.
- to enable decision support programs use at the point of clinical care: a computer-based EHR system might work with a diagnostic expert system to backing physicians´ decisions. In order to achieve optimal integration, the transference of patient information from EHR to the diagnostic expert system would need to be automated. The major barrier to do so, are the variance between the controlled vocabularies of the two systems [7, 8, 12].
- to exchange data between health information systems: the concept of Semantic Interoperability arises. We defined it as the possibility of different systems to understand the information they are processing through the use of codes of clinical terminologies.
- for epidemiological analysis: it can be used by patients, physicians, researchers, quality control and management personnel and other administrative functions like accounting, billing and coding personnel.
- for the process of codifying medical information systems actually count with vocabularies (artifacts that describe and systematize meanings of terms), with the common distinctions between terminologies (which provide standardized meanings), thesauri (which introduce semantic relations between groups of terms) and classifications (which introduce exhaustive partitions for statistical purposes). Some of them are used in an international level, while others have been defined according to local needs. (for more information, see **Table 1**)

Terminology	Collections of words or phrases, called terms, aggregated in a systematic fashion to represent the conceptual information that makes up a given knowledge domain.
Classification System	Intended for classification of clinical conditions and procedures to support statistical data analysis across the healthcare system.
Thesaurus	List of terms created from free text inputs extracted from the clinical data repository. The terms included in the thesaurus are divided into concepts (real clinical entities) and descriptions (different ways of naming these clinical entities). The thesaurus has capabilities to reject invalid terms already flagged as not appropriate for the intended use [9].
Clinical coding	Designating descriptions of diseases, injuries and procedures into numeric or alphanumeric designations. It involves the use of a EHR as the source for determining code assignment [19].
	Primary coding: clinical data are coded by the practitioner at the time of the consultation [10]. Secondary coding: a reduced number of trained persons codify the narrative text recorded by the physicians [10].
Semantic Interoperability	It refers to human interpretation of the content. There is a common comprehension among people about the meaning of the information that is being exchanged (correct interpretation is guaranteed, for this reason formal definitions of each entity, attribute, relationship, restriction and exchanged term are needed [20].

Table 1. Introduction's definition summary.

While many terminologies have been developed, no single terminology has been accepted as a universal standard for the representation of clinical concepts. By contrast, individual terminologies or components have been identified by standards organizations as candidates for specific uses [13]. The recommended terminologies include the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT); Logical Observation Identifiers Names and Codes (LOINC) and Unified Medical Language System (UMLS) [14], between many other. These will develop in the following sections [15, 16].

In the nineteenth century, the advancement of clinical pathology and technology changed the framework of classification, moving emphasis from the patient's experience to phenomena determined by physician using diagnostic procedures [17].

The diagnostic entities in medicine are changing as a consequence of expanded and revised knowledge of the functions of the human body. Techniques for differential diagnostic strategies contribute to new categories, while old labels are gradually abandoned. There is a need to acknowledge the potency of classification systems as dynamic tools for medical practice and research [17]. According to all these changes, during twentieth century, the importance of "concept orientation" in terminology construction arises. Concept orientation allows a terminology to be helpful in several situations, depicted in different languages and easily evaluated for quality [18]. This transition from the use of Classification Systems to Reference Terminology was not only a change in institution's choices, but also both of them defined their purposes, potential functions, strength and limitations.

In the following sections, we will briefly present Classifications Systems, Reference Terminology and Interface Terminology. Finally, we will present our experience developing and implementing our Terminology Server.

2. Classification systems

A classification is "a system that arranges or organizes like or related entities" [21] (for more information, see **Table 1**). Classifications provide a useful framework for a systematic representation and codification of medical concepts. Monoaxial classifications form a hierarchy of terms based on a common root. The most commonly used example of monoaxial hierarchical classifications is the International Classification of Diseases, Tenth Revision, Clinical Modification and International Classification of Diseases, Procedure Coding System (ICD-10-CM/PCS), published by the World Health Organization, represents an example of the clinical classification systems. It has been designed for providing outputs in terms of reports and statistics [5, 22, 23]. Multi-axial or multifaceted classifications combine terms belonging to different classes that themselves may be organized in a hierarchy. SNOMED is an example of this type of classification [5]. Classification systems are intended for classification of clinical conditions and procedures to support statistical data analysis across the healthcare system. They are mutually exclusive and exhaustive and they can provide standards for comparisons of health statistics at national and international levels. They have been used:

- to support other applications in healthcare including reimbursement,
- for public health reporting,
- to improve quality of care assessment,
- education,
- research,
- performance monitoring [21–23].

2.1. International classification of diseases background

Work on classification systems began in the middle of the seventeenth century with John Gaunt's refinement of the late sixteenth-century classification scheme for the London Bills of Mortality [24, 25]. International Classification of Diseases (ICD) was first adopted in Paris in 1900 [24, 25]. Architecturally, the ICD has not fundamentally changed from the sixteenth century model of the London Bills, in that each new code is added as a new row in a single list. United States did not choose to adopt ICD-10 until the end of our 25-year window, in 2015. Besides, at the time that ICD-10 was introduced, it stayed as paper book, ICD-10 was not published in electronic format [26].

In Australia, The National Center for Classifications in Health at the University of Sydney, was the first group to migrate ICD-10 into an electronic format [26].

By 2005, the WHO-FIC, an organization chartered by the World Health Organization (WHO) comprising national centers for classification around the world, created an international forum to advise on the content and evolution of WHO's Family of International Classifications (WHO-FIC) which includes ICD and the International Classification of Functioning (ICF). Currently, the WHO manages an electronic revision and update platform with WHO-FIC as a web page [26]. ICD-10-CM/PCS is an output system that was designed for general reporting purposes, public health surveillance, administrative performance monitoring, and reimbursement of healthcare services [22].

The ICD was developed to code death certificates but its use was extensive to include a large range of statistical reporting. ICD-10 has been used since the 1990s to collect mortality statistics around the world. The WHO defines coding as "the translation of diagnoses, procedures, co-morbidities and complications that occur over the course of a patient's encounter from medical terminology to an internationally coded syntax" [21]. According to this definition, ICD system has capability of being used for clinical coding and classification to enable international comparisons as regard to mortality and morbidity statistics [27].

Professional coders, who used to manually assign codes to patients' diagnoses and procedures, performed ICD-10-CM/PCS coding. Nowadays, coders use computer-assisted coding applications. These applications can facilitate accurate and efficient coding by automatically

suggesting codes based on the clinical documentation in the EHR system. Thus, ICD-10-CM/PCS coding is semi-automated at best and requires human intervention to either assign or validate selected codes [27].

2.2. Diagnosis-Related Groups (DRGs)

With the advent of capitated payments, the inevitable need of how to objectively determine severity of illness, in order to appropriately adjust capitated payments, or case mix. As outlined above, traditional disease classifications such as the ICD did not enjoy explicit severity of illness parameters; all that could be done was to infer disease severity on the basis of co-morbidity [26]. However, there may be no evidence demonstrating causality between the condition of interest and the co-morbidities. Case mix required some objective metrics and co-morbidity was it [26]. The set of measures for co-morbidity found everywhere has been the Diagnosis-Related Groups (DRGs) [28]. Since their beginning, multiple versions have continued to come out, changing architecturally combining demographic, diagnoses, and procedures into several hundred categories of care. These categories can in turn be considered to have, or not have, "complications" [26]. The 11th version, ICD-11, is now being developed through a continuous revision process, it will be finalized in 2018. For the first time, through advances in information technology, public health users, stakeholders and others interested can provide input to the beta version of ICD-11 using an online revision process. Peer-reviewed comments and input will be added through the revision period. When finalized, ICD-11 will be ready to use with EHR and information systems. WHO encourages broad participation in the 11th revision, so that the final classification meets the needs of health information users and is more comprehensive [27].

2.3. ICD: strengths and limitations

ICD's strengths include non-redundancy, meaning by this that each concept should only be expressed in one way. If two terms refer to the same concept, the sensitivity of the replies to database queries will be reduced. The ability to manage synonyms, this is important because allows the presence of authorized intermediate terms that refer to a unique term used to encode, index, and find the useful information. And finally, there are explicit relationships this refers to the types of relationships between terms in a nomenclature are clear [5].

As regard as SCS's limitations, we can name the completeness and the non-ambiguity. A full description of the medical vocabulary is very hard to achieve. About non-ambiguity, if two different types of data are stored under the same term, the specificity would be affects [5].

2.4. Others standard classification systems

Currently, many classification systems exist and are maintained by responsible agencies. Next, in **Table 2** we will briefly name and describe some of them.

ICD	The ICD is the global health information standard for mortality and morbidity statistics. ICD is increasingly used in clinical care and research to define diseases and study disease patterns, as well as manage health care, monitor outcomes and allocate resources [29].
DRG	Statistical system for classify all inpatient stay into groups for the aim of payment. The DRG classification system divides possible diagnoses into more than 20 major body systems and subdivides them into almost 500 groups. It was born for the purpose of Medicare reimbursement. In order to determine the payment, factors consider include the diagnosis involved and the resources necessary for treating the condition [30].
LOINC	Logical Observation Identifiers Names and Codes was developed to provide a definitive standard for identifying clinical information in electronic reports. Its database provides a set of universal names and ID codes for identifying laboratory and clinical test results [31]. It aims is providing a means of uniquely identifying the information elements in EHR. LOINC is remarkable for being the first completely open clinical terminology, making all content available without royalties or charges; this was driven by its creator Clem McDonald [26].
NANDA	Prior to the year 2002, "NANDA" was an acronym for the North American Nursing Diagnosis Association. In 2002, they officially became NANDA International Nursing Diagnoses Classification. They are in charge of definitions and classification of the guide to nursing diagnoses [32].

Table 2. Examples of standard classification systems.

3. Reference terminology

According to the International Standards Organization (ISO), terminologies should be formal aggregations of language-independent concepts, that concepts should be represented by one favored term and appropriate synonymous terms, and that relationships among concepts should be explicitly represented [33, 34]. The ISO specification also stated that terminologies must define their purpose and scope, quantify the extent of their domain coverage, and provide mappings to external terminologies designed for classification and to support administrative functions [33, 34]. The ISO also highlighted the value of mapping among separate terminologies designed to meet different needs. This would allow, for example, a physician to choose a concept from a clinically oriented terminology for constructing a patient's problem list and a mapped concept in an administrative classification (like ICD-9-CM) could be selected in an automated fashion for billing purposes [33, 34].

3.1. Reference terminology: a new paradigm

In 1998, J. Cimino summarized several works groups' toward defining the precise attributes of a multipurpose and shareable terminology [7, 8]. He stressed the value of "concept orientation" pending terminology construction. Concept orientation imply "*…to use concepts as basic building blocks ahead words, terms, or phrases"*. It allows a terminology to be useful in several situations, represented in different languages and easily assess for quality. For Cimino, the aim was to have a universal single clinical terminology that would cover a specialty domain's concepts completely at multiple levels of detail. Nonspecific phrases such as "not elsewhere classified" must be avoid [7, 8]. It is important to point out the need for complete and comprehensive domain coverage using non-ambiguous, non-overlapping concepts. In the absence of complete domain coverage, terminologies should integrate with other terminologies. Terminologies need to support synonymy and compositionality [35]. "High-quality vocabulary" has been defined as the vocabulary approaches completeness, is well organized and has terms whose meanings are clear [7, 8].

After Cimino's Desiderata, the difference between Terminology Systems like SNOMED CT and Classification Systems like ICD-10-CM/PCS became clearer. Both coding schemes provide the necessary data structure needed to support healthcare clinical and administrative processes. Clinical terminology systems as well as clinical classification systems were originally designed to serve different purposes and different users' requirements [27]. ICD-10 is a classification system and it was designed as an output general reporting purposes like public health surveillance, administrative monitoring, and repayments of healthcare services. For all of these reasons, a classification system can be less detailed than a clinical terminology. Contrary, SNOMED CT (**Table 3**) is a clinical terminology, it was developed to attend as a standard data infrastructure for clinical application, for these reason it requires a higher degree of specificity [36].

3.2. SNOMED CT: background

SNOMED has been used successfully on an international basis in areas such as anatomy-pathology and radiology. It has been translated into several languages. The Systematized Nomenclature of Medicine (SNOMED) nomenclature is an example of a multi-axial classification, developed by North American pathologists and extended from the Systematic Nomenclature of Pathology (SNOP) [5].

In 1965, the Systematized Nomenclature of Pathology (SNOP) was published by the College of American Pathologists (CAP) to describe morphology and anatomy.

In 1975, CAP expanded SNOP to create the Systematized Nomenclature of Medicine (SNOMED). In 1979, the most extensively adopted version of SNOMED named as SNOMED II was published. In 2000, in collaboration with Kaiser Permanente, CAP developed a new logic-based version named SNOMED RT. In the UK during twentieth century, Dr. James Read developed the Read Codes. In the end, under the National Health Service, they evolved into Clinical Terms Version 3 (CTV3). The first version of SNOMED CT, was published in January 2002, after a merge the CTV3 and SNOMED RT, performed by CAP. The merged product was called SNOMED Clinical Terms, which was shortened to SNOMED CT. SNOMED International considers SNOMED CT to be a brand name, not an acronym [37].

"Complete coverage of domain specific content"
"Use of concepts rather than terms, phrases, and words" (concept orientation)
"Concepts do not change with time, view, or use" (concept consistency)
"Concepts must evolve with change in knowledge"
"Concepts identified through nonsense identifiers" (context free identifier)
"Representation if concept context consistently from multiple hierarchies"
"Concepts have single explicit formal definitions"
"Support for multiple levels of concept detail"
"Methods, or absence of, to identify duplication, ambiguity, and synonymy"
"Synonyms uniquely identified and appropriately mapped to relevant concepts"
"Support for compositionality to create concepts at multiple levels of detail"

From Cimino et al. [7].

Table 3. Desiderata for controlled medical vocabularies in the twenty-first century.

SNOMED has been translated into several languages and successfully implemented around the world, in specialties such as anatomy-pathology and radiology. Novel development concerns the use of SNOMED as a reference terminology for health care. The next version nominated as SNOMED RT, will include data related to the causes and symptoms of diseases, treatment of patients, and the outcome of health care process [5]. SNOMED RT has the possibility to represent multiple types of hierarchies and to make the types fully explicit, after the proposed changes.

3.3. SNOMED CT as clinical reference terminology

Reference terminology was defined as a set of concepts and relationships that provide a common reference point for comparisons and aggregation of data about the entire health care process, recorded by multiple different individuals, systems, or institutions [38]. Cornet et al. defined it as "*…a system of concepts with assigned identifiers and human language terms, typically involving some kind of semantic hierarchy. Some systems may support the assignment of multiple terms, or synonyms, to a given concept…*" [26] SNOMED CT was developed to serve as a standard data infrastructure for clinical application, which requires a greater degree of specificity. A classification system can be less detailed than a clinical terminology [22]. In fact, the systems complement each other and contribute to providing quality data for different domains of the healthcare system [39]. Accordingly, both systems may be use depending on which degree of specificity is required: SNOMED is a better election to recognize unusual illness, mind while ICD-10 is consider more efficient for statistical reporting, such as collecting the top reasons of mortality and morbidity [40].

In order to accomplish "domain coverage", terminology developers have created new concepts by the utilization of two methods: pre-coordination and post-coordination. With pre-coordination, also named enumeration, is possible to model suitable levels of detail with distinct concepts, derived from real world, non-restricted usage by physicians. Generally, only clinically meaningful concepts are pre-coordinated [41]. By contrast post-coordination, also called compositionality, complex concepts can be composed from simple concepts [42]. Pre-coordination and post-coordination can complement each other, with pre-coordination providing logic and complexity and post-coordination, allowing expressivity and more complete domain coverage.

Existing terminologies that allow post-coordination are better capable to represent phrases and concepts extracted from clinical documents compare to pre-coordinated terminologies [43]. The reason is because users can both: access existing concepts and dynamically compose new concepts according to their needs, such terminologies may improve terminology domain coverage. However, even using post-coordination, it has not yet successfully modeled the entire scope of medical knowledge.

SNOMED CT provides a unified language, it may be used as a standard for communication among healthcare providers. It also highly promotes to semantic interoperability in healthcare information systems [44–46]. Its standardized logical structure and its wide acceptation make it more appropriate for high-level information exchange at national and also international levels [44–46].

SNOMED CT also includes several descriptions that can be used as an entry terminology. Finally, SNOMED CT has a standard cross mapping model; the official distribution includes data for mapping to ICD-9 (ICD—International Classification of Diseases). Additional ICD-10

cross map data has also been developed. These mappings provide the aggregate terminology features to SNOMED CT [47]. However, coding in SNOMED CT is different from conventional coding using ICD-10-CM/PCS. Coding using SNOMED CT is always automated: end users cannot view the codes assigned by the system. For this reason, software developers and EHR vendors are using SNOMED CT to help communication between different applications through a SCS. In fact, we can think of SNOMED CT as a programing language; users utilize applications that apply it without knowing what is at work in the backend [27].

3.4. SNOMED CT: strengths and limitations

SNOMED CT provides functionalities in three layers: entry terminology, reference terminology and aggregate Terminology. Between SNOMED CT's strengths, we can name the completeness, the non-ambiguity (terms must refer to only one concept), the ability to manage synonyms and finally the explicit relationships (this refers to the types of relationships between terms in a nomenclature are clear) [5].

About limitations, they include non-redundancy, meaning by this that each concept should only be expressed in one way. If two terms refer to the same concept, the sensitivity of the replies to database queries will be reduced [5]. It is also remarkable that for those cases when an institutional term cannot be represented with a standard SNOMED CT code, to create new concepts is not allow (for more information, see **Table 4**).

SNOMED	Nomenclature created by the CAP and is evolving into an international Standards Development Organization and currently regarded as the most advanced initiative in knowledge-based representations with clinical application. Each of the 300,000 terms included are defined using relationships with other terms, creating a powerful semantic network. SNOMED CT data model allows continuous extension of the nomenclature, adding new terms, always following the same compositional concept representation model, called Description Logics [48]
Terms	Collections of words or phrases, aggregated in a systematic fashion to represent the conceptual information that makes up a given knowledge domain. Terms in a terminology generally correspond to actual events or entities and to their cognitive representations in people's minds (called concepts) [14, 43]
Concept	Unit of symbolic processing in control vocabulary, a representation of a particular meaning. Concept orientation means that terms must correspond to at least one meaning and no more than one meaning. Meanings correspond to no more than one term [7, 8, 18]
Non-vagueness	Terms must correspond to at least one meaning [7, 8, 18]
Non-ambiguity	Terms must correspond to at least one unequivocal meaning and no more than one meaning, based on context. A distinction must be made between ambiguity of the meaning of a concept and ambiguity of its usage [7, 8, 18]
Non-redundancy	Meanings correspond to no more than one term [7, 8, 18]
Explicit relationships	The kind of relationships between terms in a nomenclature is not clear. Is-a, is-part-of, causes, associated-with, equivalent-to, is-in are the most usual relationships [5]
Concept orientation	Each concept in the vocabulary has a single, coherent meaning, although its meaning might vary, depending on its appearance in a context (such as a medical record). Terminologies also typically contain hierarchical organizations and other representations of linkages among concepts, such as the "is-a-type-of" relationship between "high blood pressure" and "disorder of cardiovascular system" [33, 34, 49]

Table 4. Definition summary.

4. Interface terminology

Interface terminology (IT), which has also been called colloquial terminologies, application terminologies and entry terminologies, has been defined as a systematic collection of healthcare-related phrases (terms) that supports clinicians' entry of patient-related information into computer programs [42]. But how does it happen? When health care providers type into EHR, IT links free text patient descriptors to structured, coded internal data elements used by specific clinical computer programs. Interface terminologies also facilitate display of computer stored patient information to clinician-users as simple human readable text [42]. These terminologies generally embody a rich set of flexible, user-friendly phrases displayed in the graphical or text interfaces of specific computer programs. The "entry" terminologies allow users to interact easily with concepts through common colloquial terms and synonyms. Entry terms can then map to explicitly defined concepts in a more formal terminology, such as a reference terminology, which can then define relationships among concepts [50]. EHR depend on interface terminologies for successful implementation in clinical settings because such terminologies provide the translation from clinicians' own natural language expressions into the more structured representations required by application programs [42]. Interface terminologies are crucial to foment direct categorical data entry by physicians in EHR. Historically, the efforts performed by terminology developers and the standards community, have been orientate to other kind of terminologies, like reference and administrative, instead of interface terminologies.

Between the aims of interface terminology, we can mention: to provide an institutional vocabulary for all user interfaces so they interact with known terms, including local jargon and preferences; to proportion concept lookup functions with loose lexical matches and options, to be employed for the time of data entry process of new items in a problems list or similar user interfaces. It is also important to provide short pick-lists definitions for more structured data entry in specific use templates, with a short list of valid entries and different preferred terms for the same concept in different settings. It should include the ability to accept new terms from the user, in case a concept or description is not represented and detect inappropriate terms for being too general or not valid in a subset [47].

The "usability" of an interface terminology refers to the ease with which its users can accomplish their intended tasks using the terminology. In addition, it has been demonstrated that interface terminology usability correlates with the presence of attributes that enhance efficiency of term selection and composition [51, 52]. The usability of a clinical interface terminology designed correlates with the presence of relevant insertional medical knowledge; adequacy of synonymy; a balance between pre-coordination and post-coordination; and mapping to terminologies having formal concept representations. IT enhances its usability by decreasing the number of steps required for users to find or compose the terms needed for a given task [41, 53].

Synonymy refers to the number of individual terms that can correctly represent a unique concept. Synonym types may include alternate phrases, acronyms, definitional phrases and eponyms [53]. Clinical interface terminologies are specifically designed to represent the variety

http://dx.doi.org/10.5772/intechopen.75781

of common colloquial phrases in medical discourse; rich synonymy should improve the nuance with which users can express themselves when using the terminology [53].

A very frequently asked question is why to use TS instead of only SNOMED as interface terminology? Between the reasons why we chose it, we can name:

- It is simpler for end users.
- When a single concept is not enough to define the information is possible to build a new one using post-coordination, understood as the representation of a clinical meaning using a combination of two or more SNOMED concept identifiers.
- Thesaurus allows to manage: synonyms (different descriptions related to a concept), list of valid and not recognized terms (error typing, etc), validated jargon and acronyms, list of "Not Valid" terms, thesaurus with local extension in a continuous learning process and drug composition information (commercial products) [47].
- SNOMED has pharmaceutical information as a single entity, not represented independently: Quantity of drug in the pharmaceutical presentation, measurement unit or pharmaceutical form. ut for clinical use, we need to identify single data components.

According to all the limitations mentioned before, terminology services arise.

5. Terminology services

Many definitions for terminology service exist. In previous publications, we defined as complex system of conceptual representation of medical knowledge, with relationships between concepts, with external representations of concepts in lists of standard terms (classifications) and with lexical tools that facilitate the search for terms [54].

A terminology server (TS) is a software that is composed of (**Figure 1**): a thesaurus or local interface vocabulary. This is a list of terms created from free text inputs extracted from the clinical data repository. The terms restrained in the thesaurus are split into concepts (real clinical entities) and descriptions (different ways of naming clinical entities). Thesaurus has been mapped to a reference a vocabulary, for example to SNOMED CT [9, 54]. The TS also is able to reject invalid terms before pointed out as not appropriate for the intended use [9]. The TS should also provide interactive information for refining concepts. This feature of the TS is achieved using semantic information included on SNOMED CT, navigating the subtypes/super-types hierarchy [9]. On the desiderata for TS, Chute et al. [50] attempt to articulate the functional needs of a terminology server oriented toward the clinical needs of care providers using applications in an operational environment. Between the desirable characteristics for a terminology server they included: Word Normalization, Word Completion, Target Terminology Specification, Spelling Correction, Lexical Matching, Term Completion, Semantic Locality, Term Composition, Term Decomposition (**Figure 1**).

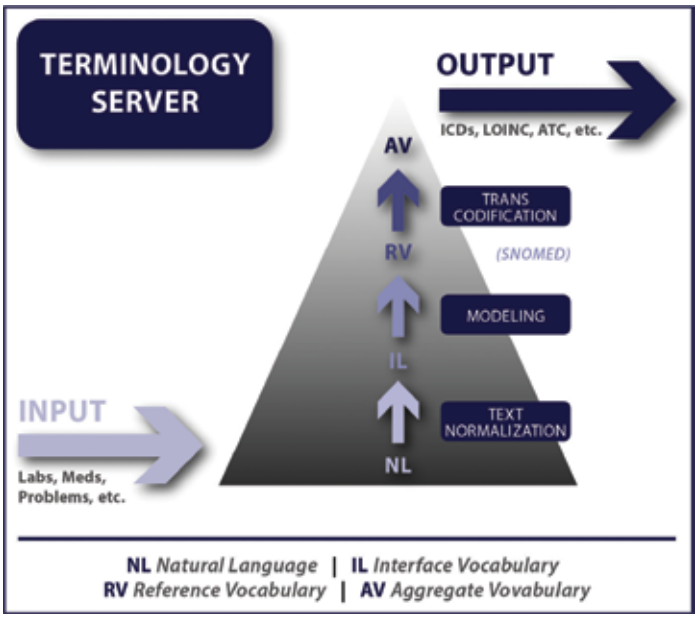

Figure 1. Schema of the functionality of the terminology server in reference to the pyramid of terminological systems.

6. Italians' hospital of Buenos Aires terminology services experience

6.1. Setting

The Hospital Italiano de Buenos Aires (HIBA) is a non-profit healthcare academic center founded in 1853, with over 2700 physicians, 2700 other health team members (including 1200 nurses) and 1800 administrative and support employees. Since 2015, it is a Joint Commission International (JCI) accredited institution. The HIBA has a network of two hospitals with 750 beds (200 for intensive care), 41 operating rooms, 800 home care beds, 25 outpatient clinics and 150 associated private practices located in Buenos Aires city and its suburban area. It has a Health Maintenance Organization (Plan de Salud) that covers more than 160,000 people and also provides health services to another 1,500,000 people who are covered by affiliated insurers. Annually, over 50,000 inpatients were admitted to its hospitals, there were 45,000 surgical procedures (50% ambulatory) and 3,000,000 outpatient visits. In addition, the HIBA is a teaching hospital, with over 30 medical residency-training programs and 34 fellowship programs. There are currently 400 residents and fellows in training. Since 1995, the HIBA runs an in-house developed health information system, which includes clinical and administrative data. Its EHR system called *Italica*, is an integrated, modular, problem-oriented and patient-centered system that works in different clinical settings (outpatient, inpatient, emergency and home care). *Italica* allows computer physician order entry for medications and medical tests, and storage and retrieval of tests results, including images through a picture archiving and communication system. In 2017, HIBA has been certified by the HIMSS as level 7 in the EHR Adoption Model, being the first hospital in Argentina and the second in Latin America reaching this stage [55]. Several health informatics standards had been implemented, including HL7, CDA Version 2, ICD-9, DRG, ICD-10, and ICPC.

6.2. Terminology server of HIBA

The terminology server of HIBA is composed of a local interface terminology (thesaurus) mapped to a reference terminology, SNOMED CT. Our main objective was to design a new terminology system, whose objectives can be related to the functions of the terminology system previously described (entry, reference and aggregate terminology) **Figure 2**.

The IT is updated every day by a team of experts, who audit, assign codes and link each new term to the SNOMED CT (reference terminology), and use the official mapping into SNOMED to another classification (like ICD 9). If SNOMED does not offer an official mapping, the team generates a manual cross-link through functionality on the terminology server [56].

6.3. Terminology server of HIBA: evolution

In 1998, the terminology work team started centralized secondary coding, where a reduced number of trained persons codify the narrative text recorded by the physicians taking care of the patients. The coding included problem list, diagnostics and procedures [10].

In 2004, we achieved 1 million of narrative text secondary coded. After this, we started an auto-codification process, through a thesaurus using interface terminology as a centralized service [56].

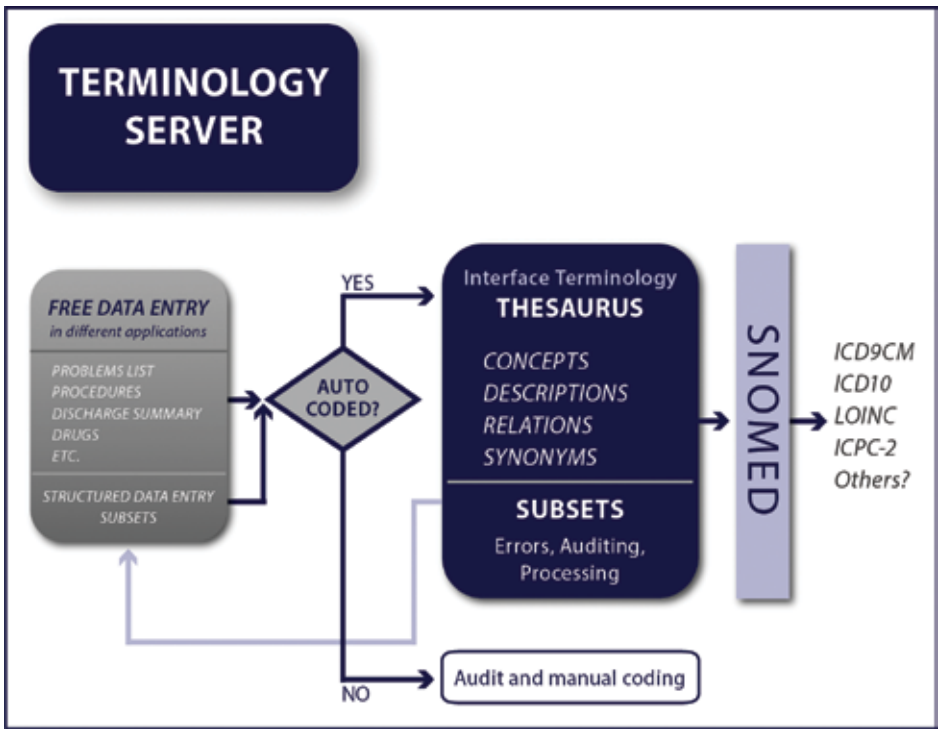

Figure 2. Schema of terminology server.

In 2010, remote Terminology Services (RTS) provided by HIBA through a transnational and interinstitutional implementation [57].

In 2011, the Startup process take places with the aim to extract the greatest amount of clinical information possible from the existing system (mostly in free text), and add this information into the new clinical data repository by coding it. To this purpose, extracted data were processed by the RTS and coded it when it was possible. This data included allergies, reason(s) for the consultation, habits, risk factors, symptoms and diagnosis entered by physicians in a free text form, and only coded diagnoses when they felt it particularly necessary. With the batch processing of these data, the RTS recognized and auto coded 11,118,760 (78.74%) texts (included valid and not valid text), and did not recognized 3,001,991 (21.26%) of the original data [57].

In 2012, we started creation of natural language processing tools and extension of terminological services to the domain of drugs, practices and procedures.

In 2014, the Department of Health Informatics of HIBA, during an effort to achieve international standards of patient health care, in the context of an accreditation process by the JCI, the hospital implemented a software tool for synchronous disambiguation in the EHR, developed in-house. Studies have shown that while the use of abbreviations helps to save time and space during documentation, its use can bring some disadvantages such as unambiguous meanings that often can confuse other healthcare providers with the consequence of causing errors in patient health care. In this sense, the JCI requires that the use of abbreviations must be controlled and documented. To this end since November 2014, an Abbreviations Regulation Committee was established in our hospital with the aim of being in charge of the management and classification of abbreviations used in historical health records. As result of this implementation, 800 abbreviations were classified as doubtful or ambiguous with a total of 400 replacement variations.

The Synchronous Self-Expanding Abbreviation System (SSAS) that detects abbreviations in a free text field. This system was user-centered design and typical abbreviations and their meanings were collected from different areas of the hospital in its construction. The abbreviations can be “unequivocal” (one meaning), “ambiguous” (more than one meaning) and “undefined” (undefined terms). SSAS detected about 4000 abbreviations (1000 univocal, 5000 Ambiguous and 2500 not defined), decreasing almost 40% in the use of abbreviations post implementation [58]. The interface vocabulary takes context parameters with terminology control such as user preferences, specialty or knowledge domain to make a decision that offers a single SNOMED CT concept. The concept-id retrieved is then used to automatically replace the abbreviation with the preferred term. The use of an interface vocabulary offers flexibility to use abbreviations with the added benefit that comes with a reference ontology [59].

6.4. The actual HIBA’s terminology web service description (Table 5)

We provide terminology services to several healthcare organizations in the countries of Argentina, Chile, and Uruguay. These include:

- a thesaurus tailored to the local needs and jargon of the professionals who interact with the EHR,
- SNOMED CT as reference standard for interoperability and to implement CDSS,

- cross maps to ICD-9, ICD-10, LOINC, ICPC-2, ATC,
- creation of different types of refsets according to the needs of the organization,
- a drug composition service modeled after the UK's dm + d model

The interface terminology is based in the use of SNOMED CT which is used as the reference terminology. In this sense SNOMED CT serves as a uniform backend representation allowing our interface terminology to adapt the local needs of the institutes we serve. SNOMED CT is the most comprehensive clinical terminology, provides a semantic network with formal structured meanings, has an extendable model, it is widely adopted as an international standard, and it was designed with EHR implementations in mind.

6.5. Our institutional entry terminology, how does it work?

The institutional entry terminology is composed of concepts and descriptions. We use SNOMED definition of these terms, where concepts represent distinct clinical meanings and descriptions are a phrase used to name a concept. Our institutional entry terminology can be divided in several subsets; examples of these are:

- Problems list terminology
- Procedures terminology
- Findings in chest radiography
- Administration routes for drugs
- State of consciousness description
- Physical examination subset
- Liver failure diagnosis

Service	Description
Inteligente prompting	Perform a preliminary search entering the first three characters.
Term recognition	Search for the text entered in the interface vocabulary and offer the alternative to improve the medical record.
Creation of a new term	Enter new term in the interface vocabulary and it is entered into the audit circuit.
List classification	Return back available classification.
Assign classifier	Valid term plus classification return back the corresponding code.
Assign DRG	From a discharge summary encoded with ICD9-CM and other metadata, returns back DRG code.
List domains	Return the domain available (problems, procedures, medications, etc.).
List domain elements	Returns back terms contained in a domain.

Table 5. Functionalities included into terminological web services provided by HIBA.

Some subsets are very large, including thousands of concepts (i.e. the problems list subset). Others are short lists (i.e. the liver failure subset). Each subset was designed in order to be used as the entry terminology in a specific scenario. Concepts are defined only once, regardless of its inclusion in more than one subset; therefore, accessing liver cirrhosis from the problems list or from the Liver failure subset brings the user to the same concept.

The process of adding concepts to the entry terminology and organizing them in subsets is manual. This is done by trained coders that were previously working with the same information in secondary coding using classifications [10]. Construction of the problems list subset was one of our biggest challenges. We decided to base our work on the historic database of our EHR with more than 2 million free text inputs since 1998. All problems list entries and discharge notes were processed to extract all different textual descriptions. We considered that these texts, entered by our own professionals in a completely unconstrained way, would be representative of the local natural language, including abbreviations and jargon. A manual depuration process, assisted by string normalization functions, led to the creation of the Problems List subset with 24,800 different concepts, with 110,000 descriptions in total.

Other subsets were created using arbitrary lists of concepts selected by the clinical terminology team with user input. New concepts or descriptions were accepted from user interfaces and stored for manual evaluation. The data model for the entry terminology was the standard SNOMED CT data model for concepts, descriptions and subsets [60]. New concepts and descriptions were added to the standard SNOMED CT distribution following official SNOMED rules for creating institutional extensions.

Since 2000, physicians at the HIBA have used an inpatient EHR for creating the discharge summary using free text. The discharge summary is a structured abstract of the hospitalization episode where data are registered for caring and management purposes. We developed and implemented a modification of the discharge summary data entry user interface that allows the selection of already coded terms from a local terminology. To achieve this we had to introduce a more restrictive user interface that requires users to select terms from an existing list. The new system should have functions that can facilitate migration from the previous unconstrained text entry model. Information contained in discharge summary is structured in several domains. This structure has the purpose of collecting all the necessary information to group episodes using DRG. In each of these fields, the physician entered free text descriptions. The previous version of the discharge summary software tried to automatically code the entered text using the terminology server. If the term did not match an existing entry in the local terminology, it was addressed to the terminology team for secondary manual codification. The terminology team reviewed all the discharge summaries, assigned ICD-9CM codes and manually grouped them into a DRG [61]. The availability of online consultation about the terminology and input terms created acceptance among users, and led us to maximize the benefits of free and structured texts [9].

6.6. Reference terminology: functions and system description

As regards as reference terminology functions, our TS allows the entry terminology should be represented in the reference terminology (SNOMED CT Spanish Language Version); new concepts can be created for institutional terms that cannot be represented with a standard SNOMED CT code. The system also provides tools to take advantage of the knowledge stored

in SNOMED CT relationships, like obtaining more refined or more general terms, and means of updating to new versions of SNOMED CT without losing information. We used SNOMED CT Spanish Language Version as the reference terminology, but it is important to note that all different language versions of SNOMED CT share the same concepts and relationships. During the translation process only, new descriptions are added. Both entry and reference terminologies were stored following the SNOMED data model, and using SNOMED tools to represent the concepts of the entry terminology. SNOMED CT defines concepts by its relationships with others, so we created new relationships as part of our SNOMED CT extension. SNOMED CT has around 300,000 concepts, but in a clinical setting, health professionals usually use very detailed expressions, adding modifiers to general concepts, like mild ankle sprain. To prevent the exponential growth of the nomenclature, SNOMED CT avoids including such level of combination with modifiers, providing the general concepts (ankle sprain), the possible modifiers (mild) and the rules to correctly relate them (using the has severity relationship).

Any new concept can be represented using this post-coordination technique, creating more detailed subtypes of existing SNOMED CT concepts. Around 33% of the concepts included in the Problems List subset could be directly mapped with existing SNOMED CT concepts; the other 77% needed the addition of one or more modifiers (post-coordination) in order to fully represent the meaning of the entry terminology concept. This rate of post-coordination was dictated by a very permissive policy allowing the use of any term requested by the users, often very specific or personalized. The total of 24,800 concepts was represented with 45,000 new relationships. In each subset, professionals usually try to enter terms that are not valid for later use. We would like the doctor to record the proper diagnosis or reason for encounter instead. In order to reject these terms and for the invalid terms administration, we tag them and add an information text so the professional understands the coding guidelines of the institution. This module provides the tools for tagging these terms and editing the information.

6.7. Aggregate terminology: functions and system description

Between the aggregate terminology functions, our TS provides output to several standard classifications: ICD-9CM (diagnosis and procedures); ICD-10 (diagnosis); ICPC-2 (diagnosis) (International Classification of Primary Care); ATC (drugs) (Anatomical Therapeutic Chemical Classification); Local billing nomenclatures; Aggregate data according to SNOMED CT hierarchies. All these functions run on a centralized software and data structure. The Terminology Server provides these functions to all existing applications in the Health Information System in the form of Web Services. A terminology maintenance software application should also be developed to administrate the institutional terminology, its relationship with SNOMED CT and the mappings.

About Aggregate terminology, the official SNOMED CT cross maps model was implemented, a multi-classification interface was created as part of the Terminology Maintenance Software to visualize, test and modify mappings from SNOMED to different classifications. An SQL algorithm was designed (Oracle SQL) to aggregate concepts according to knowledge stored in SNOMED CT relationships, like all kinds of diabetes, including diabetes complications and excluding maternal and neonatal diseases. These queries are maintained from a module in the Terminology Maintenance Software.

To code the terms in the EHR by a specific classification, the coding application requests, to select the appropriate classification. The system displays a list of classifications available and the operator must select one of them. The system then assigns the code for each term. Using this mechanism, it is possible to select the classifier ICPC-2 for the epidemiological analysis from a problem list of the outpatient EHR, ICD-9 and ICD-10 for a discharge summary in the inpatient EHR. This mapping is possible because we used the official cross-match offer by our reference terminology (SNOMED) or creates our own mapping by the specific terminology team. From a discharge summary coded in ICD-9, it may apply an assigned DRG Service to obtain the corresponding code.

6.8. Terminology maintenance software

The Terminology Maintenance Software includes the following modules:

- Entry Terminology Administration: allows the creation of new concepts, description assignment and modeling of each concept with SNOMED CT.
- Subset Administration: creation of new subsets, addition and removal of concepts from the subsets, defining hierarchies for tree interfaces.
- Pending Concepts or descriptions: all proposed new concepts or descriptions are stored in a list, waiting to be evaluated and modeled, ordered by the number of proposals.
- Cross Maps Administration: existing cross maps can be visualized, edited and tested using this module.
- Data Extraction Rules Administration: a software interface to visualize and update SNOMED based data extraction queries.

6.9. Status report

Four trained modelers are maintaining the interface terminology, modeling pending concepts or descriptions, running routine quality control checks and maintaining subsets.

We created an ad-hoc automatic process to recode all historic data in our clinical repository, using string matching algorithms; more than 2,200,000 entries were processed.

Around 85% of the original texts received a concept code of the new entry terminology, 10% of them were recognized as invalid entries: therefore, 75% were finally mapped to SNOMED CT. The coding services are used online by our ambulatory and inpatient medical record, receiving around 55,000 requests each month. The task of creating an institutional entry terminology demands a lot of work, but provides an excellent service to the users, and also isolates the terminology system from SNOMED CT changes in newer versions. Local concepts will always be valid, and in the worst case a correction of modeling against SNOMED CT would be required. We found that SNOMED CT cross maps data to ICD-9 is still not adequate for clinical use in our setting, requiring additional manual work on the maps. This may be caused by a different use of the classification in Argentina and the United States.

Our clinical data extraction process, using rules based in SNOMED CT knowledge data, is very effective; however, these rules should be revised for each new SNOMED CT version, as changes in hierarchies and models may affect its effectiveness.

http://dx.doi.org/10.5772/intechopen.75781

Further reduction of manual classification coding will require adjustments of mapping specifications and user interface changes, aimed to reduce the number of new concepts proposals and enforcing the selection of existing terms. Due to acceptability issues, we have always tried to minimize user interface constraints, thus implementation of these changes will be a slow process.

By means of a much more detailed implementation, the milestone of our new terminology system is the centralization of knowledge representation. The health information system represents uniformly the clinical data entered at any level of care in the institution.

6.10. Terminology service: experience in other settings

One of the most integrated health network of Chile, Megasalud, was using for a decade an EHR named SiapWin. In 2007, they decided to develop their own HIS allowing longitudinal care of patients treated in the network with the mentoring of the medical informatics expertise of HIBA. On behalf of this project, HIBA decided to modify the functionality of their terminology server to provide terminology services to other institutions.

In the layer of access to information, Web Services developed with JAVA, JDK 1.6 was used. The Web Services (WS) were deployed in a SUN's Glassfish application server, and the data was stored in an Oracle 11 g database. The WS were published in the Internet for the remote access of the applications of other institutions.

First implementation in a Chilean provider in 2008: they had clinical data stored and processed in the historical system: about 14 million of unique text phrases. With the terminology services, more than 11 million (78.74%) of texts were automatically codified. In 8 month about 600,000 pieces of new text were entered. About 89.64% of these new texts were successfully recognized by terminology services Nowadays, we are able to recognize above 90% in all regional implementations [57].

The clinical data stored in the legacy system of Megasalud were 14,120,751 single text phrases enabled to process by the RTS. With the batch processing of these data, the RTS recognized and auto coded 11,118,760 (78.74%) texts (included valid and not valid text), and did not recognized 3,001,991 (21.26%) of the original data. In the period between March 1 and October 1, 2009, the physicians at Megasalud entered 592,249 pieces of text in the problem-oriented EHR, 530,897 (89.64%) of them were successfully recognized in the interface terminology of Megasalud by the utilization of RTS in real time. The remainder 61,352 (10.36%) went under the audit process and manual modeling [57].

We consider great value to provide services to other institutions by our RTS. Creating and maintaining a sharable Spanish interface vocabulary database between different countries is a big task as medical Spanish is a rich vocabulary and there are different ways of naming the same clinical entities (polysemy), and different acronyms and synonyms between countries.

Published WS allow the most of the progress achieved by HIBA in the management of terminological domain. There are several services that can be used to process the text entered by a physician in their distance applications [57].

Some examples of others institutions currently consuming RTS are:

Argentina: Healthcare providers and in progress with the federal government

Chile: Healthcare providers and FONASA (National Agreement)

Uruguay: Healthcare providers and AGESIC (National Agreement)

Colombia: Healthcare providers

Actually, we are translating the thesaurus to the Portuguese, for Brazilian institutions.

Conflict of interest

The authors declare that they have no conflict of interest.

Author details

Daniel Luna*, Carlos Otero, María L. Gambarte and Julia Frangella

*Address all correspondence to: daniel.luna@hospitalitaliano.org.ar

Health Informatics Department, Hospital Italiano de Buenos Aires, Buenos Aires City, Argentina

References

[1] Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. Journal of the American Medical Informatics Association [Internet]. 2003;**10**(2):115-128 Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=150365&tool=pmcentrez&rendertype=abstract

[2] Bates DW, Boyle DL, Teich JM. Impact of computerized physician order entry on physician time. Proceedings of the Annual Symposium on Computer Application in Medical Care. 1994:996. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2247898&tool=pmcentrez&rendertype=abstract

[3] Rector AL. Clinical terminology: Why is it so hard? Methods of Information in Medicine [Internet]. Dec 1999;**38**(4-5):239-252. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10805008

[4] Shapiro JS, Bakken S, Hyun S, Melton GB, Schlegel C, Johnson SB. Document ontology: Supporting narrative documents in electronic health records. AMIA Annual Symposium Proceedings [Internet]. 2005;**2005**:684-688. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1560738&tool=pmcentrez&rendertype=abstract

[5] "Codify". Merrianm-Webster.com. Merrian-Webster, n.d. Web. 19 Apr. 2018

[6] Evans DA, Cimino JJ, Hersh WR, Huff SM, Bell DS, Group C. Toward a Medical-concept Representation Language. The Canon Group. Journal of the American Medical Informatics Association. 1994;**1**(3):207-217

[7] Cimino J. Desiderata for controlled medical vocabularies in the twenty-first century. Methods of Information in Medicine. 2012;**37**(1998):394-403

[8] Cimino JJ. Terminology tools: State of the art and practical lessons. Methods of Information in Medicine. 2001;**40**(4):298-306

[9] Navas H, Osornio AL, Baum A, Gomez A, Luna D, de Quiros FGB. Creation and evaluation of a terminology server for the interactive coding of discharge summaries. Studies in Health Technology and Informatics [Internet]. 2007;**129**(Pt 1):650-654. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17911797

[10] Luna D, De Quirós FGB, Garfi L, Soriano E, Flaherty MO. Reliability of secondary central coding of medical problems in primary care by non medical coders, using the International Classification of Primary Care (ICPC). Objetives: Materials and Methods: Results. 2001;**10**(Pt 2):94-97

[11] Hohnloser JH, Kadlec P, Puerner F. Experiments in coding clinical information: An analysis of clinicians using a computerized coding tool. Computers and Biomedical Research [Internet]. Oct 1995;**28**(5):393-401. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8612401

[12] Wong ET, Pryor TA, Huff SM, Haug PJ, Warner HR. Interfacing a stand-alone diagnostic expert system with a hospital information system. Computers and Biomedical Research [Internet]. Apr 1994;**27**(2):116-129. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0010480984710123

[13] Osenbloom STRR, Iller RA a M, Ohnson KEBJ, Lkin PELE, Rown STHB, Rosenbloom ST, et al. Interface termminologies: Facilitating direct entry of clinical data into electronic health record systems. Journal of the American Medical Informatics Association [Internet]. 2006;**13**(3):277-288. Available from: http://jamia.oxfordjournals.org/content/13/3/277.abstract\nhttp://www.ncbi.nlm.nih.gov/pubmed/16501181

[14] Humphreys BL, Lindberg DA, Schoolman HM, Barnett GO. The unified medical language system: An informatics research collaboration. Journal of the American Medical Informatics Association [Internet]. 1998;**5**(1):1-11 Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=61271&tool=pmcentrez&rendertype=abstract

[15] Humphreys BL, McCray AT, Cheh ML. Evaluating the coverage of controlled health data terminologies: Report on the results of the NLM/AHCPR large scale vocabulary test. Journal of the American Medical Informatics Association. 1997;**4**(6):484-500

[16] National Committee on Vital and Health Statistics (NCVHS). Washington, DC: Report on Uniform Data Standards for Patient Medical Record Information; 2000. p. 65. Available from: https://www.ncvhs.hhs.gov/wp-content/uploads/2014/08/hipaa000706.pdf

[17] Malterud K, Hollnagel H. The magic influence of classification systems in clinical practice. Scandinavian Journal of Primary Health Care. 1997;**15**(1):5-6

[18] Cimino JJ. High-quality, standard, controlled healthcare terminologies come of age. Methods of Information in Medicine. 2011;**50**(2):101-104

[19] AHIMA Project Offers Insights into SNOMED, ICD-9-CM Mapping Process [Internet]. Available from: http://library.ahima.org/doc?oid=60018#.Wmozc7ziaM and http://library.ahima.org/doc?oid=60018#.Wmozc7ziaM8

[20] Pan American Health Organization. eHealth in Latin America and the Caribbean: Interoperability Standards Review. Washington, DC: PAHO; 2016. Available from: http://iris.paho.org/xmlui/bitstream/handle/123456789/28189/9789275118818_eng.pdf?sequence=1&isAllowed=y

[21] Giannangelo K. Healthcare Code Sets, Clinical Terminologies, and Classification Systems. 2nd ed. Chicago: American Health Information Management Association (AHIMA); 2007

[22] Davoudi S, Dooling JA, Glondys B, Jones TD, Kadlec L, Overgaard SM, Ruben K, Wendicke A. Data quality management model (2015 update). Journal of American Health Information Management Association [Internet]. 2015;**86**(10):62-65. Available from: http://library.ahima.org/PB/DataQualityModel#.Wmo5FrziaM9

[23] Bishop J, Bronnert J, Cook J, Macksoud L, Mastronardi C, Morsch M, et al. Automated coding workflow and CAC practice guidance. Journal of American Health Information Management Association. 2010;**81**(7):51-56

[24] Chute CG. Clinical classification and terminology. Journal of the American Medical Informatics Association. 2000;**7**(3):298-303

[25] Elkin PL, Brown SH, Carter J, Bauer BA, Wahner-Roedler D, Bergstrom L, et al. Guideline and quality indicators for development, purchase and use of controlled health vocabularies. International Journal of Medical Informatics [Internet]. Dec 18, 2002;**68**(1-3):175-186. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12467801

[26] Cornet R, Chute C. Health concept and knowledge management: Twenty-five years of evolution. IMIA Yearbook [Internet]. 2016;**25**(Suppl. 1):S32-S41 Available from: http://www.schattauer.de/index.php?id=1214&doi=10.15265/IYS-2016-s037

[27] Alakrawi ZM. Clinical terminology and clinical classification systems: A critique using AHIMA's data quality management model. Perspectives in Health Information Management [Internet]. 2016. Retrieved April 2018. Available from: http://perspectives.ahima.org/clinical- terminology-and-clinical-classification-systems-a-critique

[28] Vladeck BC. Diagnostic-related groups. Journal of the American Medical Association [Internet]. Jun 25, 1982;**247**(24):3314-3315. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6806489

[29] Available from: http://www.who.int/classifications/icd/factsheet/en/

[30] DRG [Internet]. Available from: https://www.healthlawyers.org/hlresources/HealthLaw Wiki/Diagnosis-related group (DRG).aspx

[31] LOINC [Internet]. Available from: https://loinc.org/faq/basics/

[32] NANDA [Internet]. Available from: http://www.nanda.org/nanda-international-glossary-of-terms.html

[33] ISO/TS 17117:2002(E): Health Informatics—Controlled health terminology—Structure and high-level indicators: Technical Commitee ISO/TC 215, Health Informatics; 2002. p. 18. Available from: https://www.iso.org/obp/ui/#iso:std:iso:ts:17117:en

[34] International Standards Organization (ISO). TC 215 – Health Informatics. 2007. (Tc 215). Available from: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_tc_browse.htm?commid=54960

[35] Chute CG, Cohn SP, Campbell JR. A framework for comprehensive health terminology systems in the United States: Development guidelines, criteria for selection, and public policy implications. Journal of the American Medical Informatics Association. 1998;**5**(6):503-510

[36] Perspectives in Health Information Management. Fall 2014 introduction. Perspectives in Health Information Management [Internet]. 2014;**11**:1a. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25593567

[37] SNOMED CT [Internet]. Available from: https://www.snomed.org/snomed-ct/what-is-snomed-ct/history-of-snomed-ct

[38] Imel M, Campbell J. Mapping from a clinical terminology to a classification. In: AHIMA's 75th Anniversary National Convention [Internet]. 2003. Available from: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Mapping+from+a+Clinical+Terminology+to+a+Classification#0

[39] Chavis S. Two systems, one direction. For the Record [Internet]. 2013;**25**:10. Available from: http://www.fortherecordmag.com/archives/1013p10.shtml

[40] Centers for Disease Control and Prevention. International Classification of Diseases. 10th revision. Clinical Modification (ICD-10-CM). 2016. Internet [Internet]. Available from: http://www.cdc.gov/nchs/icd/icd10cm.htm

[41] Rassinoux AM, Miller RA, Baud RH, Scherrer JR. Compositional and enumerative designs for medical language representation. Proceedings of the AMIA Annual Fall Symposium [Internet]. 1997;**1997**:620-624. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2233357&tool=pmcentrez&rendertype=abstract

[42] Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. Interface terminologies: Facilitating direct entry of clinical data into electronic health record systems. Journal of the American Medical Informatics Association [Internet]. May 1, 2006;**13**(3):277-288 Available from: https://academic.oup.com/jamia/article-lookup/doi/10.1197/jamia.M1957

[43] Chute CG, Cohn SP, Campbell KE, Oliver DE, Campbell JR. The content coverage of clinical classifications. For the Computer-Based Patient Record Institute's Work Group on Codes & Structures. Journal of the American Medical Informatics Association. 1996;**3**(3):224-233

[44] Duarte J, Castro S, Santos M, Abelha A, Machado J. Improving quality of electronic health records with SNOMED. Procedia Technology [Internet]. 2014;**16**:1342-1350 Available from: http://linkinghub.elsevier.com/retrieve/pii/S2212017314003788

[45] Gøeg KR, Chen R, Højen AR, Elberg P. Content analysis of physical examination templates in electronic health records using SNOMED CT. International Journal of Medical Informatics [Internet]. 2014 Oct;**83**(10):736-749 Available from: http://linkinghub.elsevier.com/retrieve/pii/S1386505614001075

[46] Lee D, de Keizer N, Lau F, Cornet R. Literature review of SNOMED CT use. Journal of the American Medical Informatics Association [Internet]. Feb 1, 2014;**21**(e1):e11-e19. Available from: https://academic.oup.com/jamia/article-lookup/doi/10.1136/amiajnl-2013-001636

[47] Gambarte ML, Osornio AL, Martinez M, Reynoso G, Luna D, de Quiros FGB. A practical approach to advanced terminology services in health information systems. Studies in Health Technology and Informatics [Internet]. 2007;**129**(Pt 1):621-625. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17911791

[48] Spackman KA. Normal forms for description logic expressions of clinical concepts in SNOMED RT. Proceedings of the AMIA Symposium [Internet]. 2001;**2001**:627-631. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2243264&tool=pmcentrez&rendertype=abstract

[49] Hammond WE, Stead WW, Straube MJ, Jelovsek FR. Functional characteristics of a computerized medical record. Methods of Information in Medicine [Internet]. Jul 1980;**19**(3):157-162. Available from: http://www.ncbi.nlm.nih.gov/pubmed/7412563

[50] Chute CG, Elkin PL, Sherertz DD, Tuttle MS. Desiderata for a clinical terminology server. Proceedings of the AMIA Symposium [Internet]. 1999;**2**(1):42-46 Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2232621&tool=pmcentrez&rendertype=abstract

[51] Kahn CE, Wang K, Bell DS. Structured entry of radiology reports using world wide web technology. Radiographics. 1996;**16**(3):683-691

[52] The Computer-Based Patient Record [Internet]. Washington, D.C.: National Academies Press; 1997. Available from: http://www.nap.edu/catalog/5306

[53] Masarie FE, Miller RA, Bouhaddou O, Giuse NB, Warner HR. An interlingua for electronic interchange of medical information: Using frames to map between clinical vocabularies. Computers and Biomedical Research [Internet]. Aug 1991;**24**(4):379-400 Available from: http://linkinghub.elsevier.com/retrieve/pii/001048099190035U

[54] Lopez Osornio A, Gambarte ML, Otero C, Gomez A, Martinez M, Soriano E, et al. Desarrollo de un servidor de terminología clínico. 8mo Simp Informática en Salud – 34 JAIIO. 2005;**34**(April):29-43

[55] Luna D, Plazzotta F, Otero C, González Bernaldo de Quirós F, Baum A, Benítez S. Incorporación de tecnologías de la información y de las comunicaciones en el Hospital Italiano de Buenos Aires. Spanish; 2012-01. Serie: Documentos de Proyectos No.459, CEPAL; https://www.youtube.com/watch?v=5Pb_3q6TQ58&feature=youtu.be

[56] Osornio AL, Luna D, Gambarte ML, Gomez A, Reynoso G, de Quirós FGB. Creation of a local interface terminology to SNOMED CT. Studies in Health Technology and Informatics [Internet]. 2007;**129**(Pt 1):765-769. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17911820

[57] Luna D, Lopez G, Otero C, Mauro A, Casanelli CT, de Quirós FGB. Implementation of interinstitutional and transnational remote terminology services. AMIA Annual Symposium Proceedings [Internet]. 2010;**2010**:482-486. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3041368&tool=pmcentrez&rendertype=abstract

[58] Rodriguez JF, Lede DR, Pérez D, Benítez S, Luna D. Abbreviations System: Preliminary Results of a Satisfaction Study. Cbis. 2016;**2016**:1-6

[59] Williams M d PA, Perez D. Building an Snomed CT content extension with abbreviation support. Systematized Nomenclature of Medicine. SNOMED CT EXPO 2015;**2015**

[60] International Health Terminology Standards Development Organization (IHTSDO). SNOMED CT® Clinical Terms User Guide. January 2010;**2010**:99

[61] Lopez Osornio A, Luna D, Bernaldo de Quiros FG. Creación de un sistema para la codificación automática de una lista de problemas. 5to Simp Informática en Salud – 31 JAIIO [Internet]. 2002;**2002**:31. Available from: http://www.sis.org.ar/sis2002/paperssis/SIS30.pdf

Chapter 3

Multivariate-Stepwise Gaussian Classifier (MSGC): A New Classification Algorithm Tested Over Real Disease Data Sets

Alexandre Serra Barreto

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.74703

Abstract

In data mining, classification is the process of assigning one amongst previously known classes to a new observation. Mathematical algorithms are intensively used for classification. In these, a generalization is inferred from the data, so as to classify new cases, or individuals. The algorithm may misclassify an individual if the inference machine is not able to sufficiently discriminate it. Therefore, it is necessary to go further into the analysis of the information provided by the individual, until it can be sufficiently identified as belonging to a class. This chapter developed this idea for the improvement of a certain class of classifiers, using medical data sets to validate the new algorithm proposed here: The Multivariate-Stepwise Gaussian Classifier (MSGC). The results showed that MSGC is at least as competitive as the Gaussian Maximum Likelihood Classifier. MSGC attained the greatest accuracy rate in two of the data sets, and obtained identical results in the two remaining data sets. Concerning medical applications, once a classification method has been successfully validated considering a particular scope of data, the recommendable would be its use for the best diagnosis. Meanwhile, other algorithms could be tested until they proved to be effective enough to be put into practice.

Keywords: data mining, classification, algorithm, medical diagnosis and prediction

1. Introduction

Mankind has performed classification since remote years, as a part of daily life and survival. With human evolution, our motivation to classify has become more complex and wide, comprehending classification in a wide variety of fields like engineering, management, banking, marketing, psychology and medical diagnosis and prediction.

In the context of data mining, classification can be understood as the process of assigning to a new observation (sample) one among a set of previously known classes. In fact, the rapid increase in computational processing capacity, coupled with the low cost of storage, has contributed to the greater use of supervised or nonsupervised mathematical algorithms for computational classification. In these, in the learning phase, certain kind of generalization is inferred from the data, so that new cases, or individuals, can be classified by the inference machine.

It should be mentioned that in the medical field, there are several examples of researches applying successfully computational classification as an aid to the medical diagnosis. It can be referred, for instance, the research in [1], which apply a multivariate statistical analysis to explore the Dermatology Data Set (available in the UCI data repository, [2]) and construct a classifier, based only on the 12 clinical attributes, as an aid for the first medical consultation and diagnosis of erythemato-squamous dermatological diseases. The research results provide enhanced knowledge that can help to enrich dermatological diagnoses made by doctors. Also, the classifier developed using the linear discriminant analysis (LDA) obtains a high mean accuracy rate in relation to the six diseases (83.73% correct classifications). This rate means that patients have a good chance of being treated adequately, while biopsies may also be solicited to confirm diagnosis. A classification algorithm developed in [3] was tested over the Dermatology Data Set. This study reported mean accuracy rates (96.2 and 99.2% for a modified version of the algorithm). Note that it utilized all 34 features in the data set (clinical + histological attributes), which can certainly inform further the classifier, since it works knowing the biopsy results. In [4], an analysis is outlined attempting to classify the Dermatology Data Set by decision trees and employing all 34 features in the data set. The authors reported a 5.5 +/−1.46 error rate. A modified decision tree based on a genetic algorithm for attribute selection achieved a 4.2 +/−0.96 error rate. In [5], a classification algorithm is demonstrated, based on genetic algorithms that discovered comprehensible IF-THEN rules. The algorithm was submitted to all 34 features in the Dermatology Data Set and the result was 95% accuracy rate for classifications. By visiting the UCI data repository website, many other studies focusing several medical data sets are listed and can be accessed by the reader.

However, occasionally such generalization may not correctly classify an individual if the inference machine is not able to sufficiently discriminate it among the possible classes. Therefore, it is necessary to go more deeply into the analysis of the information provided by the individual being classified, until it can be sufficiently identified as belonging to a class.

This pursuing, moreover, may be analogous to the efforts made by physicians while performing their crucial diagnostics. In fact, medical theory and practice well acknowledge a basic foundation in medicine, that no two individuals are alike, either in health or illness. For this reason, more and more medical guidelines pursue this maxim, the individualities being considered in the midst of large numbers, examples being the programs of family physicians, homeopathy, psychoanalysis, encouragement of anamnesis rather than light and machine consultations and recent considerations involving slow medicine. Not to lengthen this subject too much, reference is made to the works ([6] p. 5–6, [7] p. 11–12], [8], and [9] p. 3). It could still be possible refer to a series of other initiatives that denote the search for health in its individual fullness, but what is important is that common sense says that such a foundation should also inform statistical methods and artificial intelligence applied to the classification of individuals.

http://dx.doi.org/10.5772/intechopen.74703

In this context, this chapter seeks to develop this idea for the improvement of a certain class of well-known classifiers; for this purpose, it uses real medical data sets for the validation of the algorithm proposed here.

2. Classifiers based on the assumption that the form for the underlying density function is known

In parametric classification techniques, we learn from data under the assumption that the form for the underlying density function is known. The most common procedure is to consider the normal distribution, as is the case of Gaussian Maximum Likelihood Classifier (GMLC). Suppose there are c distinct classes, given a sample vector $X^T = (x_1, x_2, \ldots, x_p)$ depicting p measurements made on the sample from p attributes, GMLC will assign to X the class h (h = 1,…, c) having the highest likelihood among the classes. GMLC assumes that the data follows the multivariate normal density function:

$$f(X|\mu_h, \Sigma_h) = \frac{1}{\sqrt{(2\pi)^p |\Sigma_h|}} \exp\left[-1/2(X-\mu_h)^T \Sigma_h^{-1} (X-\mu_h)\right]. \quad (1)$$

In this equation, μ_h is the mean vector of class h, Σ_h is the covariance matrix for class h and $|\Sigma_h|$ is the determinant of Σ_h. Usually, these parameters are not known and must be estimated from training samples. The sample mean is typically the estimate for the density mean, and the covariance matrix is usually estimated via the sample covariance matrix or the maximum likelihood covariance matrix estimate. The sample mean and the maximum likelihood covariance matrix estimates maximize the joint likelihood of training samples, which are assumed to be statistically independent [10].

The depicted above is mostly the case of some well-known classifiers like linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and regularized discriminant analysis (RDA), which are trustworthy classifiers based on GMLC computations that reach good results in several data situations. A basic difference between these three classifiers is that in the case of LDA, it is assumed that each class h comes from a normal distribution with a class-specific mean vector and a common global covariance matrix. On the other hand, QDA provides a model that assumes as many covariance matrices (Σ_h) as there are classes (h). RDA provides a kind of mix of them by means of tuning parameters (λ, γ), which provides an optimal mix of sample covariance matrix, global covariance matrix and the identity matrix, for instance, if ($\lambda = 0$ and $\gamma = 0$) RDA will represent QDA, and if ($\lambda = 1$ and $\gamma = 0$) RDA will represent LDA. It is important to note that among them there is no method considered better. For instance, in [11], it is possible to see that the performance of a classification method varies according to the database considered. The reader could refer to [12], ([13] p. 331–335) and [14] for the accessing of LDA, QDA and RDA foundations.

However, for sure, these aforementioned methods have their shortcomings. Barreto [11] lists the more commonly identified shortcomings in the field literature, such as the fact that the mean and covariance estimates are optimal only asymptotically and can produce lower

classification accuracy when the training sample is small, actually, unless many more than $p+1$ samples are available, the true covariance matrix is poorly estimated. Also, the assumption of the knowledge about the form for the underlying density function may be suspicious in most applications. Furthermore, the method involves the inversion of Σ_h estimate and in some cases, this matrix can be ill-conditioned or even singular, making matrix inversion unfeasible. In spite of the research proposing improvements, specifically concerning the covariance matrix estimation, of which RDA is a legitimate representative, these approaches remain operating under the key assumption that the form for the underlying density function is known.

Beyond these problems, this chapter wants to discuss that these methods maximize $p(X|h)\times p(h)$ to predict the class for the vector of data X, that is, $p(h|X)$. But $p(X|h)$ is calculated on the basis of the density in Eq. (1), which involves the calculation of the well-known Mahalanobis distance from the multivariate mean $[(X\text{-}\mu_h)^T\Sigma_h^{-1}(X\text{-}\mu_h)]$, which is a positive measure. Formally, the Mahalanobis distance represents a dimensionless multivariate measure of the distance between the multivariate vector X, with p dimensions, and the class mean μ_h, that also has the same p characters. The smaller the distance with respect to a specific class mean, say μ_c, the more the probability that X belongs to class c. Therefore, by the inspection of Eq. (1), it is easy to see that this mathematical density will have problems in classifying a sample that presents close values for its distances of Mahalanobis considered in relation to the means of the involved classes, a particular situation that induces misclassification errors.

The solution to fix this is to benefit both from the training set and from information proportioned by the new sample itself to be classified. Doing this, the classifier can take into account new information that will improve the overall generalization proportioned by these traditional methods. Therefore, the proposal in this chapter is to make the classification algorithm able to identify and provide treatment to the sample cases presenting close values for its Mahalanobis distances until it can reveal more clearly its actual class for the Gaussian classifier.

3. The Multivariate-Stepwise Gaussian Classifier: A new classification algorithm

What is proposed is a new classification method: The Multivariate-Stepwise Gaussian Classifier (MSGC). MSGC theoretically works on the basis of the already depicted GMLC method. Its contribution is to treat individually a sample to be classified if this sample presents close values for its Mahalanobis distances with respect to the class means involved in the classification, so that the discrimination made by the classifier is, in thesis, inconclusive. In this case, the algorithm will work employing dimensionality reduction by disregarding, one by one, in a stepwise process, the p dimensions involved in the calculation of the Mahalanobis distances until the calculated distances are dissimilar enough to give greater accuracy (likelihood) to the classification made by the method.

The key question is: what would be the best numerical dissimilarity between the distances of Mahalanobis obtained from a sample and the class means so that its classification is optimal? It can be anticipated that this response depends on the database to be focused, which will require the previous calibration of the method proposed here.

3.1. Description of the algorithm

Given c predefined classes and n sample vectors $X_i^T = (x_{i1}, x_{i2}, \ldots, x_{ip})$, $i = 1, \ldots, n$, $j = 1, \ldots, p$, depicting p measurements (dimensions) made on the sample from p attributes, x_{ij} means the jth measurement, $j = 1, \ldots, p$, for the ith sample. So, let **X** be a data matrix of type (n, p) with the measurements of data (x_{ij}) as elements, $j = 1, \ldots, p$ and $i = 1, \ldots, n$. The MSGC algorithm functions as depicted below, considering **X** as a training set [with p attributes (variables), n instances (training samples) and c classes], x_{hij} as an element of **X** belonging to a class h, $h = 1, \ldots, c$, μ_h, $h = 1, \ldots, c$, being the class mean vector for the training set, and the matrix **U** of type (v, p) as a new unknown set [with measurements of data (u_{sj}), $s = 1,\ldots, v$, $j = 1, \ldots, p$, as elements, with p attributes (variables), v cases (unknown samples) and c classes, each line of matrix **U** being a sample vector $U_s^T = (u_{s1}, u_{s2}, \ldots, u_{sp})$], these unknown sample vectors having to be classified by the algorithm. Also, consider the density $f(X|\mu_h, \Sigma_h)$ and its complete description in Eq. (1). Besides, note below that MD_{hs} refer to 'the Mahalanobis distance for $U_s^T = (u_{s1}, u_{s2}, \ldots, u_{sp})$ in relation to the class mean vector μ_h', and Δ is the 'the numerical dissimilarity between the distances of Mahalanobis (MD_{hs}, $h = 1, \ldots, c$) calculated from a sample and the class means μ_h.' The Multivariate-Stepwise Gaussian Classifier (MSGC) algorithm pseudocode is (considering $c = 2$):

---

(0) begin algorithm (initialize variables and counters, $s = 1$);

(1) while $s < v + 1$ do:

(1.1) $j = p$;

(1.2) while $j > 0$ do:

(1.2.1) calculate the mahalanobis distances MD_{hs} for $U_s^T = (u_{s1}, u_{s2}, \ldots, u_{sj})$ in relation to each class mean vector μ_h of the training data, $h = 1, 2$;

(1.2.2) calculate $f(U_s|\mu_h, \Sigma_h)$ for each class h, $h = 1, 2$;

(1.2.3) if $j = p$ then do:

$f(U_s|\mu_h, \Sigma_h) = ff_h$ for each class h, $h = 1, 2$;

end if;

(1.2.4) if MD_{1s}-$MD_{2s} < \Delta$ then do:

$j = j$-1;

else.

$j = 0$;

end do (referring to step 1.2);

(1.3) if MD_{1s}-$MD_{2s} < \Delta$.

(1.3.1) if $ff_1 > ff_2$ then do:

assign to sample U_s the class $h = 1$;

end if;

(1.3.2) if $ff_2 > ff_1$ then do:

assign to sample U_s the class $h = 2$;

end if;

else.

(1.3.3) if $f(U_s \mid \mu_1, \Sigma_1) > f(U_s \mid \mu_2, \Sigma_2)$ then do:

assign to sample U_s the class $h = 1$;

end if;

(1.3.4) if $f(U_s \mid \mu_2, \Sigma_2) > f(U_s \mid \mu_1, \Sigma_1)$ then do:

assign to sample U_s the class $h = 2$;

end if;

(1.4) $s = s + 1$;

end do (referring to step 1);

end of the algorithm.

Note that for simplicity of exposition, the above pseudocode was written for $c = 2$, but steps (1.2.4) and (1.3) can be expanded for any value for c. Another important observation to be made about the pseudocode is that if in steps (1.2.4) and (1.3) Δ is set to zero, then the new algorithm will function strictly as a GMLC.

Finally in this section, it should be added that recent literature involving classifiers which are in some way based on the GMLC method makes no mention of an algorithm that works like MSGC. See [15–28].

The Multivariate-Stepwise Gaussian Classifier (MSGC) algorithm was implemented by means of The R Program for Statistical Computing [29] (version 2.14.0).

4. Comparing MSGC with traditional GMLC method

4.1. Methodology

Some real data sets from the UCI repository [2] (available from: http://archive.ics.uci.edu/ml/datasets/) are used to compare MSGC to GMLC method.

http://dx.doi.org/10.5772/intechopen.74703

A 10-fold cross validation is widely used in the related literature like [13, 30] to present a more stable estimate of the performance of a classification method. Then, it was used here.

So to calibrate the MSGC algorithm and define the best value for Δ to be applied to each validation bloc, a previous 10-fold cross-validation was performed for each of the 10 training blocs. In the calibration process, the chosen criterion was the greatest accuracy rate reached over the 10-fold cross-validation. In this process were considered Δs starting from 0 up to 1.1, with increments of 0.1 at each iteration. If there was a tie during the process of calibration, the chosen Δ was the lower one. Thereafter, MSGC was submitted to each validation bloc (once adjusted with the best Δ regarding the corresponding training bloc and its proper process of 10-fold cross-validation).

For comparison GMLC was also implemented in the R program and applied to exactly the same blocs generated by the depicted 10-fold cross-validation process.

4.2. Presentation of data sets and comparison of classification results

Pima Indians Diabetes Data Set comprises 768 entries (8 medical and demographical attributes and a class variable), 550 of the entries classified as 0 and 268 classified as category 1. Attribute information: (1) number of times pregnant, (2) plasma glucose concentration a 2 hours in an oral glucose tolerance test, (3) diastolic blood pressure (mm Hg), (4) triceps skin fold thickness (mm), (5) 2-hour serum insulin (mu U/ml), (6) body mass index (weight in kg/(height in m)^2), (7) diabetes pedigree function, (8) age (years) and (9) class variable (0 or 1). Ten mutually exclusive folds were randomly sampled from Pima Indians Diabetes Data Set (9 validation folds including 77 entries and the tenth fold comprising 75). The key importance involved in the classification of Pima Indians Diabetes Data Set lies in the possibility of diagnosing diabetes disease, considering the numerical attributes, since class 1 is interpreted as tested positive for diabetes.

Breast Cancer Winsconsin (Original) Data Set comprises 699 entries (9 attributes and a class variable), 458 of them classified as category 2 "benign" and 241 classified as category 4 "malignant" (recoded as 0 and 1, respectively). Attribute information: (1) sample code number (id number), (2) clump thickness: 1–10, (3) uniformity of cell size: 1–10, (4) uniformity of cell shape: 1–10, (5) marginal adhesion: 1–10, (6) single epithelial cell size: 1–10, (7) bare nuclei: 1–10, (8) bland chromatin: 1–10, (9) normal nucleoli: 1–10, (10) mitoses: 1–10 and (11) class: (2 for benign and 4 for malignant). Ten mutually exclusive folds were randomly sampled from the Breast Cancer Winsconsin (Original) Data Set (9 validation folds including 69 entries and the tenth fold comprising 62). Sixteen original entries with missing data were removed. As for Breast Cancer Wisconsin (Original) Data Set, this data set can be used to predict the severity (benign or malignant) of a clump of cells in relation to the nine numerical attributes.

Haberman's Survival Data Set comprises 306 entries (three attributes and a class variable), 81 of them classified as category 2 and the remaining 225 classified as category 1 (recoded as 1 and 0, respectively). Attribute information: (1) age of patient at time of operation (numerical), (2) patient's year of operation (year-1900, numerical), (3) number of positive axillary nodes

detected (numerical) and (4) survival status (class attribute), 1 = the patient survived 5 years or longer, 2 = the patient died within 5 years. Ten mutually exclusive folds were randomly sampled from the Haberman's Survival Data Set (9 validation folds including 31 entries and the tenth fold comprising 27). The main interest in the classification task involving the Haberman's Survival Data Set would be the attempt to predict the life expectancy of patients undergoing breast cancer surgery, taking into account their age at the time of surgery and the number of axillary nodes removed.

Mammographic Mass Data Set presents discrimination of benign and malignant mammographic masses based on BI-RADS attributes and the patient's age. It comprises 961 entries of data (five attributes and a class variable). The class associated with each record is the field 'severity,' 0 or 1. Attribute information: (1) BI-RADS assessment: 1–5 (ordinal), (2) age: patient's age in years (integer), (3) shape: mass shape: round = 1, oval = 2, lobular = 3, irregular = 4 (nominal), (4) Margin: mass margin: circumscribed = 1, microlobulated = 2, obscured = 3, ill-defined = 4, spiculated = 5 (nominal), (5) Density: mass density: high = 1, iso = 2, low = 3, fat-containing = 4 (ordinal) and (6) severity: benign = 0 or malignant = 1 (binominal). A total of 131 original entries with missing data were removed. Ten mutually exclusive folds were randomly sampled from the Mammographic Mass Data Set (all of them with 83 entries). In relation to the Mammographic Mass Data Set, [2] informs that "*Mammography is the most effective method for breast cancer screening available today. However, the low positive predictive value of breast biopsy resulting from mammogram interpretation leads to approximately 70% unnecessary biopsies with benign outcomes. (…) This data set can be used to predict the severity (benign or malignant) of a mammographic mass lesion from BI-RADS attributes and the patient's age*."

To illustrate the 10-fold cross-validation process for MSGC calibration, **Table 1** summarizes the values for Δ that gave the greatest accuracy rate (%) for all the 10 training blocs. Remembering that there are two classes in all the data sets considered in the process. Afterward, these best settings for Δ (in **Table 1**) were applied to steps (1.2.4) and (1.3) in MSGC algorithm in order to classify the corresponding validation blocs.

From **Table 1**, it is possible to see that best values for Δ= 0.0 imply that MSGC optimally will work as a traditional GMLC for the Breast Cancer Winsconsin (Original) Data Set and Haberman's Survival Data Set classification.

DATA	Bloc 1	Bloc 2	Bloc 3	Bloc 4	Bloc 5	Bloc 6	Bloc 7	Bloc 8	Bloc 9	Bloc 10
PI	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.1
BR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
HB	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MA	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4

DATA SETS: PI = PIMA; BR = BREAST; HB = HABERMAN'S; MA = MAMMOGRAPHIC.

Table 1. Summary of the 10-fold cross-validation calibration process - The Δ settings giving best accuracy rate concerning training blocs.

http://dx.doi.org/10.5772/intechopen.74703

Table 2 shows synoptically the accuracy rate mean and standard error for all data sets and methods (the best results for each data sets are highlighted in bold). Both methods were proficient in classifying data and obtained relatively similar results.

From **Table 2,**we can see that MSGC attained the greatest accuracy rate in two out of four data sets (PIMA and MAMMOGRAPHIC). For HABERMAN'S and BREAST, both methods achieved identical results since for these data sets MSGC was set with $\Delta = 0.0$. It has been seen that the Mammographic Mass Data Set source [2] reports that because of the low positive predictive value of the exam, about 70% of the biopsies are actually unnecessary as they end up showing benign lesions. With the practical use of classification algorithms such as MSGC and GMLC, a favorable new situation is achieved, with levels of diagnostic accuracy above 80%. This rate means that patients can be treated adequately, while biopsies may be subsidiarily requested.

Note that accuracy rate was chosen as the criterion for comparison, but in medicine, sometimes the physician needs to know other criteria like sensitivity, specificity or precision; in this case, the data analyst should take care to also calculate them based on the algorithm results.

We also have to remark the positive aspect that these results for MSGC algorithm are transcendent. Since GMLC is the basis on which other traditional classification methods (namely RDA, QDA and LDA) are based, an improvement made in GMLC, such as this obtained through the MSGC method, will probably imply improvements in performance also for RDA, QDA and LDA. Future research shall prove this.

DATA	MSGC	GMLC
PIMA	**73.67 (2.05)**	73.41 (2.25)
BREAST	94.92 (0.75)	94.92 (0.75)
HABERMAN'S	75.10 (2.42)	75.10 (2.42)
MAMMOGRAPHIC	**80.36 (1.27)**	80.00 (1.11)

SE: Standard error for accuracy rate mean.

Table 2. Classification results for real data sets—accuracy rate mean % (SE).

5. Conclusion

A new classification algorithm is presented in this chapter: The Multivariate-Stepwise Gaussian Classifier (MSGC).

MSGC theoretically works on the basis of the Gaussian Maximum Likelihood Classifier (GMLC) method. Its contribution is to treat individually a sample to be classified if this sample presents close values for its Mahalanobis distances with respect to the class means involved in the classification, so that the discrimination made by the classifier is, in thesis, inconclusive.

In this case, MSGC will work employing dimensionality reduction by disregarding, one by one, in a stepwise process, the p dimensions involved in the calculation of the Mahalanobis distances until the calculated distances are dissimilar enough to give greater accuracy to the classification made by the base method (GMLC).

For better performance, MSGC may be previously calibrated by means of a training set. A 10-fold cross-validation process was used to calibrate the algorithm.

MSGC was applied for data classification and its performance was compared with the traditional GMLC method considering four real medical data sets available in the UCI data repository. These data represent a range of different types of data dependence structure and dimensionality. The results showed that the performance of the MSGC algorithm is at least as competitive as GMLC. MSGC attained the greatest accuracy rate in two of the data sets (PIMA and MAMMOGRAPHIC). For HABERMAN'S and BREAST data sets, both methods achieved identical results. It was concluded that MSGC can be used as an effective classification tool in a wide range of data sets.

The presented results for the MSGC algorithm are transcendent. Since GMLC is the basis on which other traditional classification methods (namely RDA, QDA and LDA) are based, an improvement made in GMLC, such as this obtained through the MSGC algorithm, will probably imply improvements in performance also for RDA, QDA and LDA.

After reaching the conclusions, an additional discussion arises. With the emergence of the big data, as a robust successor to data mining emerged from the exponential development of computers and storage media since the 1990s, it has been a tendency to think of the intensive use of multiple algorithms simultaneously, in supervised or nonsupervised approaches, to analyze data and discover patterns. This certainly makes sense, as it has already been mentioned in this chapter that there is no one classification method or algorithm better than another. Beyond to a greater robustness or scalability of some methods over others, what concrete exists is a dependence of the results against the target database.

Therefore, in this context, and considering a matter as important as the medical clinic, once a classification method has been tested and successfully validated considering a particular scope of data, the most recommended would be its use for the best diagnosis. Meanwhile, if possible, already known or new algorithms could be tested for various diseases and symptoms data until they proved to be robust and effective enough to be put into medical practice.

Finally, it is important to remark that mathematical classifier serves as an aid to the crucial medical diagnosis made by the physician.

Acknowledgements

The author would like to thank the UCI Machine Learning Repository and the data donors for putting real data sets at the disposal of the scientific community and would also like to thank The R Foundation for Statistical Computing and its contributors for developing and making the R program available to the public. The Breast Cancer Wisconsin (original) Data Set was

obtained from the University of Wisconsin Hospitals, Madison, from Dr. William H. Wolberg, and the author would like to thank him.

Nomenclature

GMLC	Gaussian Maximum Likelihood Classifier
LDA	linear discriminant analysis
MSGC	Multivariate-Stepwise Gaussian Classifier
QDA	quadratic discriminant analysis
RDA	regularized discriminant analysis

Author details

Alexandre Serra Barreto

Address all correspondence to: alexsbdr@gmail.com

Ministry of Finance (Brazil), Brasília-DF, Brazil

References

[1] Barreto AS. Multivariate statistical analysis for dermatological disease diagnosis. In: Proceedings of the IEEE International Conference On Biomedical And Health Informatics (BHI '2014); 1-4 June 2014; Valencia. New York: IEEE; 2014. p. 500-504

[2] Dua D, Karra Taniskidou E. UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Science. 2017. http://archive.ics.uci.edu/ml

[3] Demiroz G, Guvenir HA, Ilter N. Learning differential diagnosis of Eryhemato-squamous diseases using voting feature intervals. Artificial Intelligence in Medicine. 1998;**13**(3):147-165. DOI: 10.1016/S0933-3657(98)00028-1

[4] Pappa GL, Freitas AA, Kaestner CAA. Attribute selection with a multi-objective genetic algorithm. In: Proceedings of 16th Brazilian Symposium on Artificial Intelligence (SBIA); 11–14 November, 2002; Recife: Springer; 2002. p. 280-290

[5] Fidelis MV, Lopes HS, Freitas AA. Discovering comprehensible classification rules with a genetic algorithm. In: Proceedings of the 2000 Congress on Evolutionary Computation (CEC00). 16-19 July, 2000; La Jolla. New York: IEEE; 2000. p. 805-810

[6] Grandgeorge D. The Spirit of Homeopathic Medicines. Berkeley: North Atlantic Books; 1998

[7] Freud S. The Essentials of Psichoanalisys. London: Penguin Books; 1986

[8] Ali N, Khuwaja A, Kausar S, Patients NK. Evaluations of family practice care and attributes of a good family physician. Radcliffe Publishing, Quality in Primary Care. 2012;**20**: 375-383

[9] Rakel R. Family physician. In: Rakel R, Rakel D, editors. Textbook of Family Medicine. 8th ed. Philadelphia, PA: Elsevier Saunders; 2011. p. 1-14. DOI: 10.1016/B978-1-4377-1160-8.10001-6.ch1

[10] Hoffbeck JP, Landgrebe DA. Covariance matrix estimation and classification with limited training data. Pattern Analysis and Machine Intelligence. 1994;**18**(7):763-767. DOI: 10.1109/34.506799

[11] Barreto AS. Weighted correlation matrix similarity: A new classification algorithm. In: Proceedings of the IADIS European Conference on Data Mining 2012: Part of the IADIS Multi Conference on Computer Science and Information Systems 2012 (IADIS DATA MINING 2012); 17–23 July 2012; Lisbon. Lisbon: IADIS; 2012. p. 79-90. ISBN: 978-972-8939-69-4

[12] Friedman JH. Regularized discriminant analysis. Journal of the American Statistical Association. 1989;**84**(405):165-175. DOI: 10.1080/01621459.1989.10478752

[13] Venables WN, Ripley BD. Modern Applied Statistics with S. 4th ed. New York: Springer; 2002. 495 p. ISBN: 0-387-95457-0

[14] Rpubs. Classification: Linear Discriminant Analysis [Internet]. 2014. Available from: https://rpubs.com/ryankelly/LDA-QDA [Accessed: 2017-12-24]

[15] Skolidis G, Sanguinetti G. Bayesian multitask classification with Gaussian process priors. IEEE Transactions on Neural Networks. 2011;**22**(12):2011-2021. DOI: 10.1109/TNN.2011.2168568

[16] Maugis C, Celeux G, Martin-Magniette M-L. Variable selection in model-based discriminant analysis. Journal of Multivariate Analysis. 2011;**102**(10):1374-1387. DOI: 10.1016/j.jmva.2011.05.004

[17] Hyun-Chul Kim Z, Ghahramani Z. Bayesian Gaussian process classification with the EM-EP algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006; **28**(12):1948-1959. DOI: 10.1109/TPAMI.2006.238

[18] Opper M, Winther O. Gaussian processes for classification: Mean-field algorithms. Neural Computation. 2000;**12**(11):2655-2684. DOI: 10.1162/089976600300014881

[19] Ye J, Janardan R, Li Q, Park H. Feature reduction via generalized uncorrelated linear discriminant analysis. IEEE Transactions on Knowledge and Data Engineering. 2006; **18**(10):1312-1321. DOI: 10.1109/TKDE.2006.160

[20] Guo P, Jia Y, Lyu MR. A study of regularized Gaussian classifier in high-dimension small sample set case based on MDL principle with application to spectrum recognition. Pattern Recognition. 2008;**41**:2842-2854. DOI: 10.1016/j.patcog.2008.02.004

[21] Halbe Z, Aladjem M. Regularized mixture discriminant analysis. Pattern Recognition Letters. 2007;**28**:2104-2115. DOI: 10.1016/j.patrec.2007.06.009

[22] Ji S, Ye J. Kernel uncorrelated and regularized discriminant analysis: A theoretical and computational study. IEEE Transactions on Knowledge and Data Engineering. 2008; **20**(10):1311-1321. DOI: 10.1109/TKDE.2008.57

[23] Licheng J et al. An organizational coevolutionary algorithm for classification. IEEE Transactions on Evolutionary Computation. 2006;**10**(1):67-80

[24] Liu J, Chen S, Tan X. A study on three linear discriminant analysis based methods in small sample size problem. Pattern Recognition. 2008;**41**:102-116. DOI: 10.1016/j.patcog.2007.06.001

[25] Lu H, Plataniotis KN, Venetsanopoulos AN. Uncorrelated multilinear discriminant analysis with regularization and aggregation for tensor object recognition. IEEE Transactions on Neural Networks. 2009;**20**(1):103-123. DOI: 10.1109/TNN.2008.2004625

[26] Lu H, Plataniotis KN, Venetsanopoulos AN. Regularized discriminant analysis for the small sample size problem in face recognition. Pattern Recognition Letters. 2003;**24**:3079-3087. DOI 10.1016/j.patcog.2007.06.001

[27] Peng J, Zhang P, Riedel N. Discriminant learning analysis. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics. 2008;**38**(6):1614-1625. DOI: 10.1109/TSMCB.2008.2002852

[28] Xu P, Brock G, Parrish R. Modified linear discriminant analysis approaches for classification of high-dimensional microarray data. Computational Statistics & Data Analysis. 2009;**53**:1674-1687. DOI: 10.1016/j.csda.2008.02.005

[29] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN: 3-900051-07-0. URL: http://www.R-project.org/.2011

[30] Jiao L, Liu J, Zhong W. An organizational coevolutionary algorithm for classification. IEEE Transactions on Evolutionary Computation. 2006;**10**(1):67-80. DOI: 10.1109/TEVC.2005.856068

Chapter 4

Moving towards Sustainable Electronic Health Applications

Sahr Wali, Karim Keshavjee and Catherine Demers

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75040

Abstract

Electronic healthcare applications, both web-based and mobile health (mHealth) provide new modalities for chronic disease. These tools allow patients to track their symptoms and help them manage their condition. The sustainability of these tools is often not considered during their development. To ensure these applications can be adopted and sustainable, where policy differs amongst states and provinces, we must present the benefits of our findings to highlight the justification for its development. For technology to be sustainable it has to utilize infrastructure that is secure, stable and to be agile so that it can be deployed quickly with minimal interruption to patients, family members and healthcare professionals.

Keywords: sustainability, self-care, eHealth, mHealth, technology, co-design

1. Introduction

Within the healthcare industry, innovation remains to be the leading force in the quest to balance health care quality and costcontainment [1]. Mobile health (mHealth) applications are one of the fastest growing segments in drive for innovation in the health care sector. With the rising use of mobile phones, mHealth applications (apps) provide individuals with a simple and accessible way to manage their health at the tip of their fingers [1].

Unfortunately, many mHealth interventions continue to be developed without the consideration of long term sustainability, which has left many apps with vast potential but nowhere to move forward. This is one of the growing problems with health app development, where in spite of the advances made with technology, apps fail to be used due to the methodological

challenges associated with designing for sustainability [2]. In this chapter, we focus on addressing the main issues app publishers face during the design process. We then outline the key components that should be included to assure the sustainability of an electronic health app.

We define the issues with innovation by three main components that include (1) end-user usability, (2) clinician and informal caregiver (spouse, children, friend) input and (3) impact of agencies outside development. Many electronic health apps fail to consider these major factors in its design, which in turn is often what limits the sustainability of its use. We believe that these three factors are essential as it evaluates the health apps design according to the user, their main members of care and finally the environment it is used. If the health app being designed does not simplify or improve the current model of care, there is no incentive for its use. Instead, the benefits associated with the app will be overshadowed by its complications or drawbacks.

This leads to our section in the chapter on designing for sustainability. We start by signifying the importance of putting the end-user first and then introduce the information system research framework that help identify user needs, design preferences and potential barriers to increase health app adoption. This is followed by the next stage of designing for sustainability, where we outline the steps to get all the potential players on board in support for the new innovation. We highlight the inevitable resistance to change that will occur, and explain the concept of 'behavioral intention' to use a technology and how this will help improve health app sustainability.

Finally, for the purpose of long-term sustainability, we expand on preparing for the expected and unexpected, by evaluating change management plans and regulations in place during health app design. Towards the end of the chapter, we develop a market and feasibility analysis framework for the adoption and scalability of the health app on a national scale. This allows us to ensure all key factors have been addressed; leaving the app design and efficacy to become unquestionable.

2. Issues with innovation

Over the past decade, there have been a number of advancements within the healthcare industry, yet there is still a strong resistance present towards the implementation of health innovation [3]. The lack of certainty in the interventions independent sustainability is one of the leading factors responsible for this resistance [4]. Electronic health apps may hold great promise for better health tracking [5], providing education [6], changing and enforcing health behaviors [7], and monitoring treatment adherence [8], however despite these benefits, they are still not being used [9]. This can be attributed due to nine key design barriers that are outlined in Chindalo et al. literature review [9] (**Table 1**).

These barriers have created a stigma around stakeholders investing in health app development. Currently, the perceived return on investment (ROI) with health apps remains low, as the issues with innovation remain high. The beneficial impact of health apps may seem promising, but from a sustainability standpoint, they fail to address the underlying question that is 'will these benefits outweigh the cost of its development?'

Barrier	Explanation
1. Apps provide information conflicting from what is received from clinicians	Patients/end-users are less likely to use an app when it conflicts with information from their clinician. They will not feel confident in the content provided or its functionality
2. Language used too complex for end-user health literacy level	Patients/end-users often have lower levels of health literacy. They require technology to be adapted to their needs or the app will not be used sufficiently
3. Manual data input required	Treatment regimens are already perceived as complex by patients/end-users. Manual data input further complicates this process, as it is exhaustive and error-prone
4. Information provided has no value/meaningless data	If the app information has no beneficiary tie to the patient (e.g. cannot order diagnostic testing or prescribe medications) then the content becomes useless
5. Daily app use not required	Health apps aiming to help patients/end-users with their treatment regimen should be used daily and in accordance with their prescribed treatment. It the health app does not require daily use, this can reduce treatment adherence as the patient will not get into the habit of using it
6. Lack of incentives to use	Any source of change is viewed as burdensome, thus, if no incentive (cost savings or social approval) for a patient/end-user to utilize the tool is present, they are less likely to use it
7. Data collected not valued by clinician	If the data collected brings forward information important for the patient/end-users care, the clinician would be more likely to promote its use. However, if there is no functional value for the data, both clinicians and end-users/patients will not use the tool
8. No way for physicians to use data collected	Health apps may collect large amounts of data, but if they cannot visualize or analyze the meaning behind the data, it comes useless
9. No way to integrate app data into electronic medical record (EMR) for analysis or follow-up	If the data collected cannot be combined with previous medical information the context required for analysis will be lost, leaving the data to be meaningless

Table 1. Design barriers associated with decreased health app usage (adapted from [9]).

To ensure that resources being spent on an application are adequately being used and the above barriers are addressed, the needs, wants and expectations of the health apps primary stakeholders should be evaluated [10]. However, it is this lack of stakeholder consideration within app design that builds the three prime issues with innovation, which we describe below [1, 9].

1. **Poor end-user usability: who are you designing for?**

   The overall hype of innovation, and mHealth solutions, has led developers into a cycle where app ideas centered on addressing patient challenges seem to forget about the patient once in development [11]. Consequently, this lack of end-user engagement has led health app usage to fall to 2% amongst patients at hospitals in the US [11]. The low percentage for health app usage may seem surprising, but when a tool does not suit the needs or capabilities of the end-user, the percentage becomes less surprising and more understandable.

   Findings reveal that patients with chronic disease, such as heart failure and diabetes, have positive attitude towards using mobile technology if they are simple and effective [12].

However, the key issue here is that app developers seem to show greater motivation by the cleverness of the technology rather than the improvements in health outcomes, which often results in complicating the apps functionality [13]. Thus, in the eyes of the developer the app may seem effective, but they do not consider that the individual they are designing it for will not have the same understanding. As a result, apps will not meet user needs or capabilities, which in turn leads to the development of the first six barriers highlighted in **Table 1**. This poor product usability can be attributed due to the lack of end-user involvement or input during the development process [11]. Some would argue the most successful health apps are those that address real-life challenges in the context that the patient lives. Therefore, to assure the sustainability of a health application the question the developer must ask is not 'does it solve a problem', but rather 'does it help the patient directly'. If an application is in any way a burden, or adds more effort into their treatment, it will not be used.

2. **Lack of clinician and informal caregiver input: What are you designing it for and how will it improve clinical outcomes?**

   The primary objective for a health application is to improve clinical outcomes and reduce the level of work required, clinicians and informal caregivers (spouse, children, friend) play a pivotal role in establishing the criteria for these improvements [14]. Clinicians provide the complete medical background surrounding patient care as well as clinical workflow operations, while informal caregivers allow developers to have a magnified look into the day-to day challenges that prevent adherence and worsen symptoms [3, 14]. Both key members of care contribute substantively in increasing adherence, improving self-care, quality of life and outcomes for patients [14]. However, the reality is, numerous apps are/have been developed with either none or some clinician and/or caregiver feedback, but the inclusion of both are pivotal to assure sustainability.

   As every tool must have an objective for its development, the use of clinician and caregiver input provides developers with the necessary content to build their objective around. The lack of clinician and caregiver feedback in current health apps limits app efficacy and is responsible for design barriers seven and eight in **Table 1**. Without the consideration of both the *physiological* and *social* factors provided by clinicians and caregivers, health apps will continue to be designed around the question 'does it a solve a problem', and developers will inevitably fall short in the effectiveness of their app design [10].

3. **Fail to consider impact of agencies outside development: does it effect current operations?**

   Aside from the lack of usability considerations, other factors including, government regulations and organization operations, are also commonly neglected. The development and implementation of healthcare innovations is bounded by a set of regulations that must be followed [15]. These regulations are set in place as a standard to prevent public health risk and improve patient safety. In the United States, the Food and Drug Administration (FDA) issued a draft for the regulation of mobile medical applications [16]. However, as many of the standards currently in place are set to regulate medical devices, a large group of apps still do not fall within the categories for regulation. This leaves them to be generated without regulatory precaution or guidance, which in turn promotes the development of less effective and integrative health apps [17]. For example, one key element that would

increase app use and sustainability would the use of data integration amongst EMRs and the respective health app. This would allow the data collected to be combined with current medical information, which would potentially improve clinical outcomes and future diagnosis [9]. However, as this is not a required component within health app development, it has become a barrier, instead of a benefit, for health app usage (**Table 1**).

Nevertheless, one area of regulation all health apps must adhere to involves privacy and security of personal health information. Ensuring that the introduction of a health app does not threaten the privacy of the data obtained is one of the pillars for a sustainable health intervention. However, as many health apps are very development-focused these pivotal components are not acknowledged early on. This results in complicating the design process and restraining development in the later stage [10]. Similarly, as many workflow operations may be changed with the introduction of an electronic health app, failing to consider the necessary requirements early in design process leads implementation to become more disruptive than beneficial [18]. An effective health app design would identify the key organizational barriers and resistance points that may occur prior to actual implementation. Thus, the problems associated with security and workflow operations are built upon the underlying issue that they are not considered until the start of implementation.

Furthermore, electronic health apps may have undeniable potential to improve health outcomes in a cost-effective manner, but the underlying issues with innovation are preventing their potential from being fully utilized [15]. Stakeholders may have different objectives for the outcome of a health app, but regardless, the app designer must still address their needs regarding usability, content, safety, clinical and cost-effectiveness accordingly. There is already a resistance for innovation in healthcare; therefore, to build a model for health app sustainability, we outline a series of frameworks to minimize the occurrence of these issues.

3. Designing for sustainability

Prior to designing an electronic health application or any innovative health tool, we must consider the who, what, where, when and how of the intervention. Who will be using it? What will it do? Where and when will it be used? Lastly, how will it work? These questions allow the designer of the tool to recognize its primary stakeholders, the risks and costs of innovation and whether it will work with current operations in place [10]. When designing for sustainability the goal should be to bridge a practical solution for a prominent issue. Therefore, addressing these common questions has framed our guideline for moving towards sustainable electronic health applications. In this section, we start by identifying the end-user and their needs, followed by an outline to creating a user-centered health app, and finally end with the steps required to gain support for app implementation.

3.1. Putting the end-user first

Whether older adults with heart failure or adolescents with diabetes will use a health application, identifying the end-user and their needs at the start of the design process is pivotal for

the next steps towards development [19, 20]. Nevertheless, although the importance behind end-user evaluation has been signified, various studies confirm the lack of health apps available suiting their needs and capabilities [1, 21–23]. In the Delphi study, a literature review was conducted overviewing the determinants of innovation in health care organizations [24]. Their results indicated that many innovation studies failed to adjust their strategies according to feedback obtained, or that the data on the determinants was insignificant as it came from non-users [24]. In this case, the study highlights that it is not enough to simply obtain random feedback, but we must obtain useful input and apply it into the design [24].

Often times, the benefits of the end product overshadow the content required for adequate usability, leaving both the app developer and the end-user at a disadvantage. For example, by simplifying intervention processes and health education it is estimated that this will improve clinical outcomes. However, what is not considered is that the sustainability of these benefits will only be seen when the app is user-friendly and end-users can independently utilize it with confidence [25, 26]. To get to this stage, developers must recognize what components intended users need, so it becomes both easy to use and useful. Thus, to accomplish this, a user-centered framework has been developed which we summarize below.

3.1.1. Information system research (ISR) framework

As many electronic health interventions are designed according to the current healthcare system processes, this limits their impact potential compared to those that involve end-user input [27]. The ISR framework uses three research cycles, (1) relevance cycle, (2) design cycle and (3) rigor cycle, to identify user needs, design preferences as well as any barriers that will prevent the uptake or sustained use of the app [27, 28].

In the *Relevance Cycle*, developers or researchers seek out to understand the end-user in the context of their environment [19]. It is the environment that shapes the specificities behind the arising problems, the purpose of this cycle is to provide the requirements for the health app, as well as set of criteria for users to evaluate its functionality [28, 29]. Thus, to meet the goals of this cycle focus group style sessions with intended stakeholders and end-users are commonly used [27, 29]. By the end of this cycle, we should be able to answer the question, 'Does this app improve the user's environment, and how?' [28].

The heart of development occurs during the *Design Cycle*, as the content from the relevance cycle is used to build the health app and evaluate it accordingly [28]. This cycle continues in an iterative manner, where a series of designs will be generated and evaluated against the respective user requirements, until all key components are addressed. The design of the app can move relatively quickly, however, it is the continual evaluation and feedback for refinements that challenges developers [29]. Nevertheless, end-users often stop using apps that do not immediately engage them, so by repeatedly conducting prototype testing with key stakeholders, this increases the expected usability and sustainability of the end product.

Finally, the *Rigor Cycle* is the background check of the ISR framework [29]. It reviews and evaluates the current knowledge base present within the desired applications domain [27, 28]. This enhances the degree of innovation for the health apps design. In many cases, this cycle is conducted after the relevance cycle to increase the overall effectiveness of the apps design [27] (**Figure 1**).

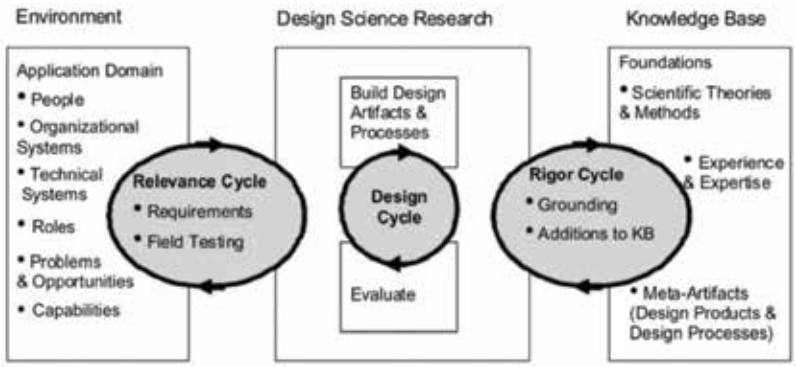

Figure 1. The ISR framework divided into three design science research cycles, (1) relevance cycle, (2) design cycle and (3) rigor cycle [28].

3.1.2. Creating a user-centered design: ISR and end-user co-design

The foundation of a user-centered design is centralized on three major components, (1) understanding how the device will be used, (2) curating information relevant to the end-user and (3) framing the tool in the user's environmental context and lifestyle [4]. The ISR framework allows developers to assess the needs of the end-user while evaluating current interventions in place [28]. However, the co-design method moves one step further by using a participatory approach where end-users and primary stakeholders work together on all aspects of the health apps development [30, 31]. By using the ISR framework in parallel with the co-design method, we believe this iterative process will lead towards a more effective user-centered end product over the long-term [29, 31].

Many electronic health app interventions fail to engage users in the design and usability stages [31]. In a systematic review of co-designed mHealth interventions, studies included patients in the development stages, but none assessed the intervention's effectiveness afterwards [31]. Conversely, in another study, users evaluated the interventions usability, but were not involved in its design [32]. The lack of fluidity between mHealth development and user input reduces end-user empowerment and overall app usability. The healthcare system is already burdened with various pre-mature innovation investments that have fallen short in its beneficial return. Therefore, from a sustainability standpoint, by using the ISR framework, this will allow all factors surrounding the end-user and the current knowledge base to be covered, whereas the co-design aspect will be pivotal to assure its usability.

3.2. How to get everyone on board

One of the greatest obstacles towards developing sustainable electronic health interventions involves getting primary stakeholders in support for its development and implementation [3]. This challenge has been shaped due to the three paradoxes of innovation [33]. First, the uptake of the dubious and rejection of the good. The explosion of electronic health apps created a consumer fad where a number 'breakthrough' apps left individuals in regret and stakeholders reluctant to invest again. Second, the wisdom and failings of democracy. Working with professional groups can be effective to ensure implementation of a new technology,

however, solely relying on their cooperation results in killing the product before it is even complete. Third, health systems are not able to keep up. Innovation results in causing change in an organization, but this creates challenges that innovators are often not prepared for and results in causing more disruption than improvement [33].

In order to move past these challenges we must be address the following questions:

1. What evidence is there that it will improve outcomes and how will it effect current operations?

2. Will any additional support be needed before it can be introduced?

3. How should it be monitored during introduction?

The first question allows us to determine whether the electronic health app will be worth the investment. The second and third questions are key for its sustainability, as it recognizes components pivotal for a smooth implementation procedure [33]. Breaking down the barriers built by failed innovative interventions may be difficult, but it is beyond worthwhile to develop an effective health app. Answering these questions will be essential when developing a plan to obtain stakeholder support, thus we further discuss the specific steps to break down the resistance and prepare for the change below.

3.2.1. Battling the resistance to change

With any type of change there is an inevitable build-up of resistance that is formed. This resistance is derived from the fear of failure, similar to the first paradox of innovation; executives and end-users do not want to waste their time with another unbeneficial intervention [3, 33]. With this in mind, assimilating the idea of putting a new intervention into practice will be uphill road to climb. Nevertheless, two models described below help shape the key factors and steps involved towards achieving this goal.

3.2.1.1. Technology acceptance model (TAM)

The TAM was developed to drive the use of new technology and increase its acceptance by assessing the end-users *perceived ease of use* (*PEOU*) and *perceived usefulness* (*PU*) [33, 34]. This model suggests that by clarifying that a new source of innovation will reduce the amount of effort required (PEOU) and enhance performance it will be more likely to be accepted amongst end-users and other key stakeholders [35]. Thus, in the context of health care, executives, clinicians and patients will find an electronic health app more useful and user-friendly if they are familiar with the technology [3, 34].

With this in mind, when designing an electronic health app, developers must understand what the stakeholders needs, wants and expectations are. Once this is discovered, we can adequately highlight how the health app will benefit each of them specifically. Finally, when a foundation of acceptance for the app has been established, appropriate training protocols should be instilled to prevent any former resistance from re-establishing (**Figure 2**).

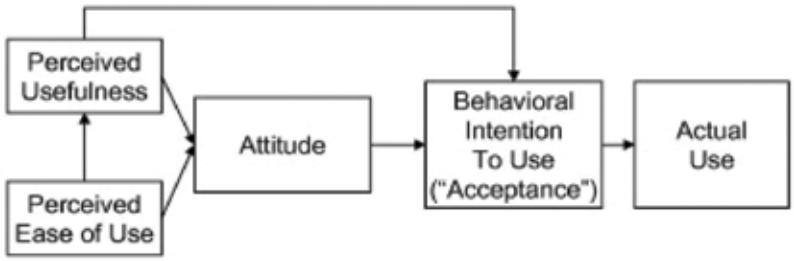

Figure 2. The technology acceptance model assessing the end-users perceived ease of use and perceived usefulness of technology to determine their behavioral intention and actual usage potential [34].

3.2.1.2. Diffusion of innovation (DOI) theory

The DOI theory is used to increase the adoption of technology [3]. This is one of the oldest theories, yet it remains to be continually used during innovative design. The DOI theory states that an organization will consider a technology to be innovative if it is perceived as new and relevant. It proposes that four main elements contribute to the diffusion of an idea, (1) the idea itself, (2) communication channels, (3) time and (4) a social system [36]. Similar to the TAM, the DOI theory suggests highlighting the perceived advantage and relevance of the innovations development. Therefore, in the context of electronic health apps, the benefit the of the app should be communicated amongst various influencers, and then stakeholder support must be established before the app can be readily adopted. As health apps must be widely adopted amongst all stakeholders and end-users before it can become self-sustaining, tackling this challenge will be key for the apps long-term success.

3.2.1.3. Presenting the benefits: results to support longevity

In both the TAM and the DOI theory, the key message to help obtain stakeholder support was to simply present why the app is beneficial for them. Why should they care about what we are developing? The real question developers should ask is, 'How will it help them?' This leads into one of the key components towards breaking down the barrier of stakeholder resistance and moving towards designing a sustainable health app. To adequately present the benefits of an application we must have the appropriate evidence to support our claim.

In many cases, a health app may be the first of its kind or an idea may be an advancement of a previous intervention. Regardless of whether a pilot study has been conducted to support the benefits of its use or its benefits have yet to be evaluated, the success of the product can be supported by answering the same question mentioned above, 'How will it help them'. We must outline what the problem currently is and why the development of this health app will help address it. It is not enough to state that a problem exists; it is the reasoning behind the solution that highlights the justification for its development.

Moreover, it is important to present the evidence supporting the benefits of the health app, but we must also present this support in the context of each stakeholder. Depending on the type of app being designed the stakeholders will differ, but to increase each claims value, we must understand the factors that will influence health app acceptance and evaluate them accordingly.

3.2.1.4. Testing health app acceptance: unified theory of acceptance and use of technology (UTAUT)

After assessing stakeholder needs and presenting the advantages associated with the health app, one of the next steps towards sustainability is to evaluate whether the proposed solution will be accepted amongst various users [37]. The UTAUT is a technology acceptance model commonly used to predict a user's behavioral intention to use a technology [37]. This model is based on four key components, [38].

1. **Performance expectancy:** providing an incentive to use a technology is key to ensure its acceptance. Performance expectancy is defined as the extent that the technology will benefit the end-user in completing a certain task. It is expected that increasing the health app's beneficial value, this will increase users behavioral intention. Therefore, to present these benefits, performance expectancy is constructed by four main evaluative criteria: (1) perceived usefulness: how much they will believe the technology will improve their performance, (2) extrinsic motivation: what other valued outcomes (money, fame) they will receive from its use, (3) job fit: how suitable technology is to increase performance and (4) relative advantage: benefit of new technology compared to what it will cost [39].
2. **Effort expectancy:** ease of use is a critical component of technology acceptance. Effort expectancy refers to how easy or difficult it is viewed to use the technology. Past technology acceptance models, such as TAM and DOI, have signified how applications that are simpler to use are more often accepted [40]. Thus, to reduce effort expectancy and increase acceptance rates, health apps should be less complex and instead easier to use [39].
3. **Social influence:** in many cases, the decision making process is influenced by specific individuals or the social norm. Social influence is the degree a user perceives that key individuals (family, friends) believe the use of the technology is important. This can be caused by informational influence where information from other people impact a decision or it can be normative influence where a user conforms their decision according to what is defined as 'acceptable' according to a certain group or situation. Regardless, social influence can be an ultimate determinate regarding the overall acceptance and usage of a health app [39].
4. **Facilitating conditions:** to ensure technology acceptance, users must feel its implementation is both feasible and realistic. Facilitating conditions is the degree individuals believe that existing organization and technical infrastructure is present to support its usage. These conditions can vary depending on a health app's objective, but regardless, they have a significant impact on its adoption and usage [40].

These four concepts have a direct influence on the behavioral intention to use a health app. Age, gender, experience and voluntariness are also associated with indirectly influencing behavioral intention and technology usage, as they moderate the four UTAUT component

relationships. Thus, by incorporating the UTAUT within the health app design process, this will allow us to predict users' intention to adopt the app in an organizational context [38]. Nevertheless, as majority of health apps are focused on the users setting and the challenges they face, we recommend the usage of the more updated UTAUT2.

Compared to the original UTAUT, the extended UTAUT2 has shown to improve the variance in behavioral intention and technology use [38]. The UTAUT2 builds on the core UTAUT principles around extrinsic motivation, and adds three other components to improve the prediction of behavioral intention, which we describe below.

1. **Hedonic motivation:** hedonic motivation is defined as how enjoyable the technology is to use. It is used to analyze the emotional and psychological aspect of the technology, which is often overlooked by most evaluative models. The functionality of a health app will only go so far in influencing technology acceptance; it is the user experience that ultimately determines its long-term use and sustainability. Therefore, by evaluating the users internal satisfaction, this will result in improving technology usage [38, 39].
2. **Price value:** the price value determines whether the benefit of the technology is greater than its monetary cost. If the price value is high then individuals feel the benefit of use is greater than the cost of investment. However, this is not always the case and it is pivotal to assess this aspect of health app design to ensure its sustainability [38].
3. **Habit:** habit is the degree individuals automatically perform tasks due to learnt behaviors [38]. This construct was added to the UTAUT2 as it helps assess whether user-activity will be sustained. Often times, there is a fall out in health app usage due to the tasks becoming burdensome. Thus, if we were able to make these tasks more like a reflex than an extra step, this would reduce the effort required and improve technology acceptance (**Figure 3**).

With the addition of these factors to the UTAUT2, this helps tailor the health app evaluation to users in their context, which in turn helps improve its overall acceptance (**Table 2**). In many cases, it is these factors that prevent the sustainability of a health app. App publishers are focused on eliciting a behavior change or improving clinical outcomes, so they tend forget about the individual in their context. To reduce the resistance to change and move towards acceptance, it is the responsibility of the app publisher to ensure that the tool they are introducing is not only effective, but also it is 'fun', affordable, easy and relatable, or else it will simply not be used. Therefore, we believe by using the UTAUT2 principles during health app design, this will allow for the development of a more effective product, and will lead into a smoother transition phase during implementation.

3.2.2. Preparing for the expected and the unexpected

The implementation of any new source of innovation will come with many challenges that are both expected and unexpected, however, being prepared for both is what ensures optimal sustainability. We highlight below two of the major areas that result in impeding health app implementation, which is (1) regulations and (2) change management. Both are essential to incorporate during the design plan of a health app, we describe the detailed steps we recommend to tackle both areas effectively.

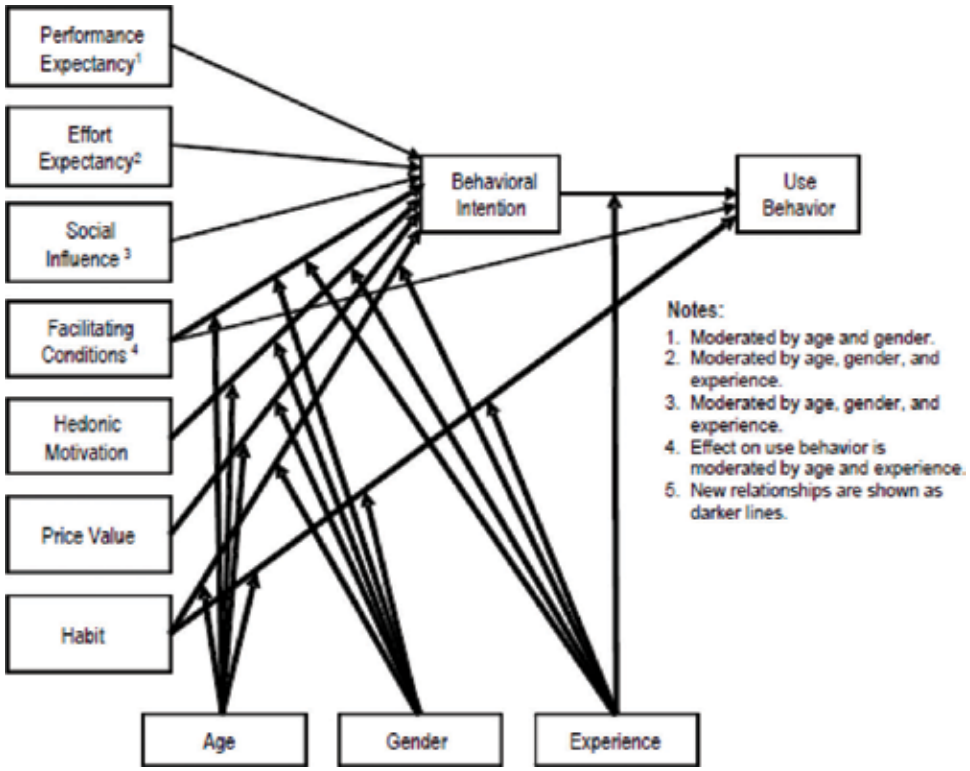

Figure 3. The UTAUT2 model with addition of hedonic motivation motivation, price value and habit to determine behavioral intention and technology use. All factors contribute in influencing behavioral intention to use except for 'facilitating conditions' and 'habit' which directing effect use behavior [38].

3.2.2.1. What regulations?

When implementing any new intervention into the health care industry there are a series of regulations that must be reviewed [15]. The FDA and Health Canada have issued a set of restrictions when developing an electronic application used as a medical device; however, most health apps do not fall under this category [17]. Nevertheless, one aspect of regulation all health apps must oblige to involves privacy and security. As many health apps are mobile phone based, this creates a challenging situation where more data can be obtained but data privacy is not secure. Policymakers are still in the works of establishing specific criteria required for patient safety, however we have listed a series of components that should be included within the health apps design to protect data integrity and prevent any unexpected threats.

(1) Data Sharing and Consent Management—Who can share my data

All data shared must have consent, as well as meet the Health Insurance Portability and Accountability Act (HIPAA) standards for data sharing [17].

(2) Access Control and Authentication—Who can access data

To assure that only the approved individuals can access the data, an authentication procedure should be implemented. With most health data, data encryption and a respective login passcode are usually required.

(3) Confidentiality and Anonymity—Who knows it's my data

Depending on the level of consent and the data obtained, most personal data should remain confidential and possibly anonymous if used for public-health purposes [17].

Higher degrees of data security protocols can be implemented into health app designs, however, we believe that by incorporating these three aspects of privacy and security, this will make the app more desirable for both its end-users and its respective stakeholders during development. Ultimately, being prepared with the proper security measures gives stakeholders the confidence in the product, and will ease the process of change management.

			1	2	3	4	5
Performance expectancy	PE1	I find the health app useful in my daily life					
	PE2	Using the health app increases my chances of achieving things that are important to me					
	PE3	Using the health app helps me accomplish things more quickly					
	PE4	Using the health app increases my productivity					
Effort expectancy	EE1	Learning how to use the health app is easy for me.					
	EE2	My interaction with the health app is clear and understandable					
	EE3	I find the health app easy to use					
	EE4	It is easy for me to become skillful at using the health app					
Social influence	SI1	People who are important to me think that I should use the health app					
	SI2	People who influence my behavior think that I should use the health app					
	SI3	People whose opinions that I value prefer that I use the game					
Facilitating conditions	FC1	I have the resources necessary to use the health app					
	FC2	I have the knowledge necessary to use the health app					
	FC3	The health app is compatible with other technologies I use					
	FC4	I can get help from others when I have difficulties using the health app					

			1	2	3	4	5
Hedonistic motivation	HM1	Using the health app is fun					
	HM2	Using the health app is enjoyable					
	HM3	Using the health app is very entertaining					
Price value	PV1	The health app is reasonably priced					
	PV2	The health app is a good value for the money					
	PV3	At the current price, the health app provides a good value					
Habit	HT1	The use of the health app has become a habit for me					
	HT2	I am addicted to using the health app					
	HT3	I must use the health app					
	HT4	Using the health app has become natural to me					
Behavioral intention	BI1	I intend to continue using the health app in the future					
	BI2	I will always try to use the health app in my daily life					
	BI3	I plan to continue to use the health app frequently					

Table 2. UTAUT2 questionnaire used to evaluate health app acceptance amongst end-users.

3.2.2.2. Change management is key for smooth sailing

With the introduction of any new intervention this will result in causing changes in workflow that in some cases may be disruptive. These challenges are expected, but to assure the implementation process runs smoothly, a set of change management plans can be pre-developed [41]. To develop a proper strategy, we must consider three primary levels of change management.

(1) Individual change management: how people experience the change and what their needs are to successfully make the transition.

(2) Organizational/initiative change management: what are the primary groups that will directly be impacted and what changes will need to be completed respectively.

(3) Enterprise change management capability: this is the overall organizational approach to managing change. It usually involves executive discussion, and reflects the organizations capability to allow and embrace change. This level is key as top-down support has a direct relationship on how a change will be perceived at the lower levels.

All three levels of change management can be addressed through the three-phase change management process (**Figure 4**) [41]. Phase 1, prepare for the change, we must determine who will be impacted by the change and what level of support we will need to smoothly move forward. During this phase, it will be key to understand all the challenges that will be in play, as

Figure 4. The change management process indicated by three phases (1) preparing for change, (2) managing change and (3) reinforcing change [41].

we will need instill that the perceived usefulness will remain to be stronger. Phase 2, manage the change, focuses on supporting the individuals impacted by the change. With respect to the implementation of health apps, this phase would be heavily focused on the end-users and what additional training that may be required to increase its perceived ease of use and prevent resistance from re-establishing. Phase 3, reinforcing change, evaluates the current status of the intervention to identify any issues and address them accordingly. This phase is key for the long-term sustainability of the health app as it ensures the change is maintained and provides evidence to support its benefits [41]. By following this three-phase change management plan, the health app design process will move more efficiently and will lead to a higher adoption rate.

4. Ensuring adoption and scalability

Moving from end-user usability and primary stakeholder needs, app publishers must also consider the underlying factors for national adoption. State or province specific regulations, needs, and resources available are likely to differ across a country. Thus, for optimal scalability, technology should be agile enough to utilize the infrastructure that is in place without causing disruption.

For instance, in Canada, policies regarding home-care differ between provinces, leaving some provinces with funding and others with none. When introducing a health app with similar objectives to home-care, developers must consider how the app can be used independently and concurrently, without losing its value. To ensure long-term adoption and scalability, the ideal health app should be designed to seamlessly coincide with the current practices that are in place. We describe below what evaluative steps we recommend to maximize national health app potential.

4.1. Long-term sustainability: where will it work?

We have described components should be included to please various stakeholders and ensure a smooth transition process. However, true sustainability stems from its capability to be seamlessly used in multiple settings. To accomplish this, we must conduct a market analysis to understand what interventions and regulations are currently in place, followed by a feasibility assessment to determine if those factors will jeopardize its implementation on a national scale [2, 42].

Depending on the health app being developed, different factors contributing to the market will be evaluated. However, before evaluating any components, the objective of the health app must be distinguished. This will narrow the scope of market factors that we will need to consider. We will then need to determine whether the health app will be an improvement of a current intervention or if it will stand alone in its functionality. Once this has been decided, we can readily evaluate what market factors are at play by answering a set of questions as outlined in **Figure 5**. Moving forward, depending on the con-current usage of the health app the steps may differ. Nonetheless, both ends of the evaluation will move in the same direction towards evaluating app feasibility, by determining what regulations/policies are in place, how cost-effective the app will be and most importantly if there is any funding available to support its implementation (**Figure 5**). This evaluative framework regarding the market forces and feasibility will determine whether the app will be sustainable across national regulations, or if it requires substantial changes in its design construct.

In the health app industry, it is common for enthusiasm regarding innovation to overshadow the drive to tackle sustainability, let alone feasibility. In this chapter, we focused on the highlighting the importance beyond designing sustainable electronic health applications. We started by addressing what barriers regarding sustainability were present and outlined what steps were needed to avoid them. The importance of identifying the needs, wants and expectations of the health apps primary stakeholders were also signified, as we understand that it is not only

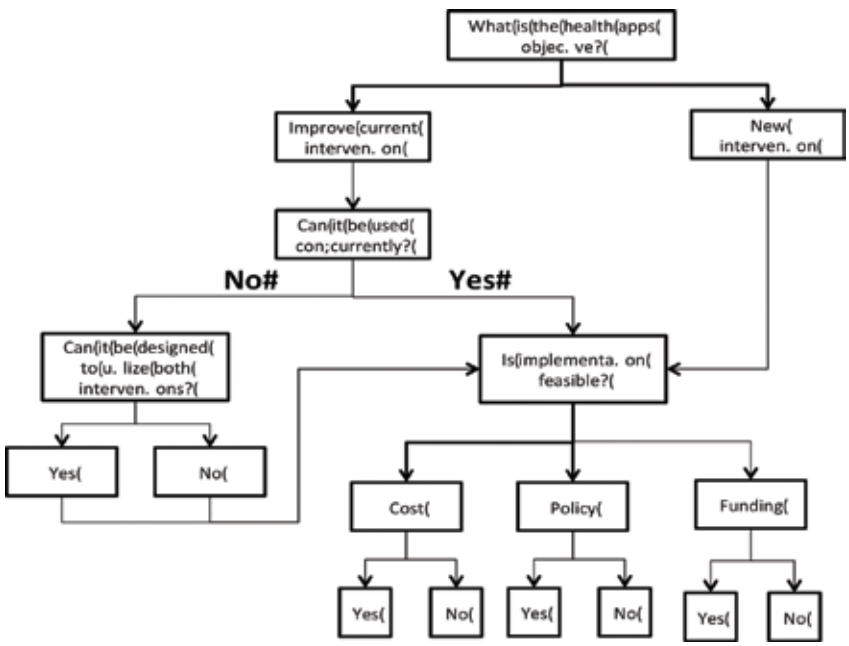

Figure 5. Market evaluation and feasibility analysis framework to help determine national sustainability of the intervention. The analysis begins by highlighting the health apps objective and determining whether it will improve current processes or introduce a new one. This leads to evaluating how the health app will function with current practices in place and if it will be suitable moving forward. In this framework, each block represents a form of analysis that is conducted, which leads to the final block that assesses overall feasibility and/or potential health app re-design/modifications needed.

important from a usability standpoint, but it shapes whether it will be sustainable across the country. Designing for sustainability may be a tiresome process, but if executed properly, the end results will bring more value than anticipated.

Author details

Sahr Wali[1], Karim Keshavjee[2] and Catherine Demers[1]*

*Address all correspondence to: demers@hhsc.ca

1 McMaster University, Hamilton, Canada

2 InfoClin Inc, Toronto, Canada

References

[1] Research2guidance. App Developer Economics 2015. 2015. Available from: http://research-2guidance.com/r2g/r2g--App-Developer-Economics-2015.pdf [Accessed: November 27, 2017]

[2] Proctor E, Luke D, Calhoun A, McMillen C, Brownson R, McCrary S, Padek M. Sustainability of evidence-based healthcare: Research agenda, methodological advances, and infrastructure support. Implementation Science. 2015;**10**(88):1-13. DOI: 10.1186/s13012-015-0274-5

[3] Thakur R, Hsu SHY, Fontenot G. Innovation in healthcare: Issues and future trends. Journal of Business Research. 2012;**65**:562-569

[4] Marshall D, Demers C, O'Brien B, Guyatt G. Economic evaluation. In: Dicenso A, Guyatt G, Ciliska D, editors. Evidence-Based Nursing: A Guide to Clinical Practice. St Louis: Elsevier Mosby; 2005. pp. 298-317

[5] Abril EP. Tracking myself: Assessing the contribution of mobile technologies for self-trackers of weight, diet, or exercise. Journal of Health Communication. 2016;**21**(6):638-646

[6] Hartin PJ, Nugent CD, McClean SI, Cleland I, Tschanz JT, Clark CJ, Norton MC. The empowering role of mobile apps in behavior change interventions: The gray matters randomized controlled trial. JMIR mHealth uHealth. 2016;**4**(3):e93

[7] Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: Pilot randomized controlled trial. Journal of Medical Internet Research. 2013;**15**(4):e32

[8] Cook KA, Modena BD, Simon RA. Improvement in asthma control using a minimally burdensome and proactive smartphone application. The Journal of Allergy and Clinical Immunology. In Practice. 2016;**4**(4):730-737

[9] Chindalo P, Keshavjee K, Karim A, Brahmbhatt R, Saha N. Health apps by design: A reference architecture for mobile engagement. International Journal of Handheld Computing Research. 2016;**7**(2):34-43

[10] Omachonu VK. Innovation in healthcare delivery systems: A conceptual framework. Public Sector Innovation Journal. 2010;**15**(1):1-20

[11] Patient Empowerment Network. Designing with the patient in mind. 2016. Retrieved from: https://powerfulpatients.org/2016/01/26/designing-with-the-patient-in-mind/

[12] Kuerbis A, Mulliken A, Muench F, Moore AA, Gardner D. Older adults and mobile technology: Factors that enhance and inhibit utilization in the context of behavioral health. Mental Health and Addiction Research. 2017;**2**(2):1-11

[13] Patient Empowerment Network. Are we ready for mobile health. 2016. Retrieved from: https://powerfulpatients.org/2016/01/04/are-we-ready-for-mobile-health/

[14] Kotz D, Gunter CA, Kumar S, Weiner JP. Privacy and security in mobile health: A research agenda. IEEE Computer. 2016;**49**(6):22-30

[15] Faulkner A, Kent J. Innovation and regulation in human implant technologies: Developing comparative approaches. Social Science & Medicine. 2001;**53**(7):895-913

[16] Hollis C, Moriss R, Martin J, Amani S, Cotton R, Denis M, Lewis S. Technological innovations in mental healthcare: Harnessing the digital revolution. British Journal of Psychiatry. 2015;**206**(4):263-265

[17] Barton AJ. The regulation of mobile health applications. BMC Medicine. 2012;**10**(46):1-4

[18] Avgar AC, Litwin AS, Pronovost PJ. Drivers and barriers in health IT adoption. Applied Clinical Informatics. 2012;**3**:488-500

[19] Aidemark J, Asskenas L, Nygardh A, Stromberg A. User involvement in the co-design of self-care support systems for heart failure patients. Procedia Computer Science. 2015;**64**: 118-124

[20] McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, Cafazzo J. mHealth consumer apps: The case for user-centered design. Horizons: Technology and Design. 2012;**46**:49-56

[21] Brahmbhatt R, Niakan S, Saha N, Tewari A, Pirani A, Keshavjee N, Mugambi D, Alavi N, Keshavjee K. Diabetes mHealth apps: Can they be effective. Studies in Health Technology and Informatics. 2017;**234**:49-53

[22] Martinez-Perez B, Torre-Diez I, Lopez-Coronado M, Herreros-Gonzalex J. Mobile apps in cardiology: Review. JMIR mhealth uhealth. 2013;**1**(2):e15

[23] Creber RMM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the mobile application rating scale (MARS). JMIR mHealth uHealth. 2016;**4**(2):e74

[24] Fleuren M, Wiefferink K, Paulussen T. Determinants of innovation within health care organizations: Literature review and Delphi study. International Journal for Quality in Health Care. 2004;**16**(2):107-123

[25] Gandhi S, Chen S, Hong L, Sun K, Gong E, Li C, Yang LL, Schwalm JD. Effect of mobile health interventions on secondary prevention of cardiovascular disease: Systematic review and meta-analysis. Canadian Journal of Cardiology. 2017;**33**:210-231

[26] Bonoto BC, de Araujo VE, de Lemos LL, Godman B, Bennie M, Diniz LM, Junior AA. Efficacy of mobile apps to support the care of patients with analysis of randomized controlled trials. JMIR mHealth uHealth. 2017;**5**(3):e4

[27] Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, Gelaude D, Patterson Mosley D, Travers J. A user-centered model for designing consumer mobile health (mHealth) applications (apps). Journal of Biomedical Informatics. 2016;**60**:243-251

[28] Hevner ARA. Three cycle view of design science research. Scandinavian Journal of Information Systems. 2007;**19**(2):87-92

[29] Cronholm S, Gobel H. Evaluation of the information systems research framework: Empirical evidence from a design science research project. The Electronic Journal Information Systems Evaluation. 2016;**19**(3):158-168

[30] Donetto S, Pierri P, Tsianakas V, Robert G. Experience-based co-design and healthcare improvement: Realising participatory design in the public sector. International Journal for All Aspects of Design. 2015;**18**(2):227-248

[31] Eyles H, Jull A, Dobson R, Firestone R, Whittaker R, Te Morenga L, Goodwin D, Ni Mhurch C. Co-design of mhealth delivered interventions: A systematic review to assess key methods and processes. Current Nutrition Report. 2016;**5**(30):160-167

[32] Lloyd T, Buck H, Foy A, Black S, Pinter A, Pogash R, Eismann B, Balaban E, Chan J, Kunselman A, Smyth J, Boehmer J. The Penn state heart assistant: A pilot study of a web-based intervention to improve self-care of heart failure patients. Health Informatics Journal. 2017:1-12. Web

[33] Dixon-Woods M, Amalberti R, Goodman S, Bergman B, Glasziou P. Problems and promises of innovation: Why healthcare needs to rethink its love/hate relationship with the new. BMJ Quality and Safety. 2011;**20**(1):47-51

[34] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly. 1989;**13**(3):319-340

[35] Holden RJ, Karsh BT. The technology acceptance model: Its past and its future in health care. Journal of Biomedical Informatics. 2010;**43**(1):159-172

[36] Rogers EM. Diffusion of Innovations. 5th ed. New York, NY: Free Press; August 2003. pp. xv-xxi

[37] Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS Quarterly. 2003;**27**(3):425-478

[38] Venkatesh V, Thong JYL, Xu X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly. 2012;**36**(1):157-178

[39] Huang CY, Kao YS. UTAUT2 based predictions of factors influencing the technology acceptance of phablets by DNP. Mathematical Problems in Engineering. 2015;**2015**:1-23

[40] Chang A. UTAUT and UTAUT2: A review and agenda for future research. The Winners. 2012; **13**(2):106-114

[41] Prosci. What is change management. 2018. Retrieved from: https://www.prosci.com/change-management/what-is-change-management

[42] Swerissen H, Crisp BR. The sustainability of health promotion interventions for different levels of social organization. Health Promotion International. 2004;**19**(1):123-130

Chapter 5

The Practice of Medicine in the Age of Information Technology

Mark Dominik Alscher and Nico Schmidt

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75482

Abstract

Regarding the practice of medicine, we have to face the chances and challenges of all aspects of e-Health; however, the term "digitalization" is broader and spanning all aspects. However, the digitalization of medicine offers solutions for pressing problem. We know the factors that lead to excellence in medicine. Without the right amount of experiences based on a solid ground of knowledge, no excellence is achievable. The problem, nowadays, is that due to restriction of working hours, to the goals of life ("life-work-balance") and the restrictions of Generation Y, almost no education in medicine is spanning the needed 10,000 h experiences in practical medicine for excellence. Therefore, we will see the fading of medical excellence, if we could not establish other systems. A solution can be searched in decision-support systems. However, a requirement before is the need of a digitalization of all health data. We surely do not have enough evidences for all aspects of the practice of medicine, the intuition is fading away and therefore, we have to look around for other solutions. Big data generated by the digitalization of all health data could be the problem solver. In combination, IT will help to improve the quality of care.

Keywords: quality, practice of medicine, digitalization, health care, intuition, big data, randomized controlled trial (RCT)

1. Introduction

Nowadays we found a lot of changes of the frame works for all professions. The terms "digitalization," "Internet of Things," "disruption," and "big data" cover some aspects of these changes on different hierarchical levels. Regarding the practice of medicine, we have to face the chances and challenges of all aspects of e-Health; however, the term "digitalization" is broader and spanning all aspects [1]. In the following chapter, I try to highlight some aspects especially in the face of practical medicine.

2. Excellence in the practice of medicine

We know the factors that lead to excellence in medicine. Without the right amount of experiences based on a solid ground of knowledge, no excellence is achievable. However, without knowledge and without the ability for processing the experiences, excellence cannot be found [2]. Therefore, experiences alone are not the key to excellence [3]. It is the combination of genius, knowledge base, and experiences. Simon and Chase have found for master chess player that 10 years of practice are necessary [4]. In that time period, around 50,000 different patterns are stored that are essential for the intuitive part of the game. Ericsson was able to extend these findings to musicians and physicians [2]. For excellence, you must be worked in practice for about 10,000 h.

However, let us have a closer look to excellence in the practice of medicine. Colleagues were asked what makes the difference regarding excellence in their colleagues [5]. They gave four factors:

1. Extensive practical experiences.
2. Master in taking the medical history from patients.
3. Precise and critical integration of all information into the process of diagnosis reasoning.
4. Continuous learning of clinical practice.

In internal medicine, the process of diagnostic reasoning is key to excellence [6]. This process can be divided into two parts:

System 1: Intuition

System 2: Analysis

Both have different properties (**Table 1**). For the doctor in the practice of health care, due to time pressure and the big number of patients and problems a typical doctor has to treat, handle, and solve in short time, the system 1 (intuition) is the only practical way, that is in most part confined to the amount of experiences made before. However, from time to time,

System 1: Intuition	System 2: Analysis
Experimental-induction	Hypothesis-deduction
Rational limitations	Rational unlimited
Heuristic	Normative
Pattern recognition	Robust decisions
Modular (“hard-wired”) decisions	Critical-logical thinking
Guidance by pattern recognition	Decision trees
Gut feeling	Logical reasoning

Table 1. Medical decision-making after Croskerry [6].

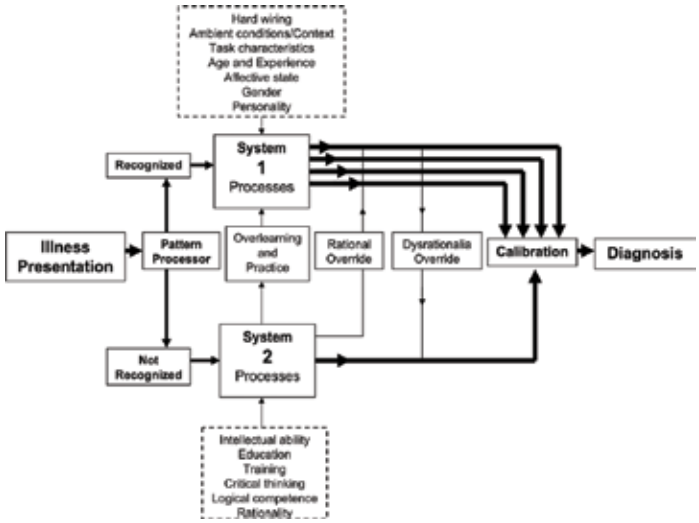

Figure 1. Combination of System 1 and System 2 in diagnostic reasoning [6].

the doctor will face a problem, he cannot solve by intuition, then he has to go to system 2 (analysis), that is time-consuming. Ideal would be an automatic adjustment of both systems in all decision-making during the process of diagnostic reasoning (**Figure 1**).

The problem nowadays is, that due to restriction of working hours, to the goals of life ("life-work-balance") and the restrictions of Generation Y, almost no education in medicine is spanning the mentioned 10,000 h experiences in practical medicine. Therefore, we will see the fading of medical excellence, if we could not establish other systems to replace system 1. The United States was leading in guarantee excellence in medicine by education since the days of Flexner [7]. The base of the excellence, however, was the precise and holistic learning of the medical knowledge base [8]. This is under pressure [9–11].

A solution can be searched in decision-support systems. However, before we need a digitalization of all health data. Electronic patient records are the key to accomplish that task [12]. Since we do not have the holistic solution, interfaces, and standards for that interfaces are highly needed [13, 14]. The analysis of data from the health system, often called "big data," is confined on a solution to those issues. Algorithms should give help in a world of overwhelming information load for the doctor and should release him from the pain of long working hours. This is the promise of big data from the viewpoint of the practitioner.

3. Big data versus randomized controlled trials (RCT)

Measurement of quality in diagnostic reasoning and decision-making is the evidence-based decision, based on precise evidences generated, at best, in randomized controlled trial (RCTs) [15]. RCTs revolutionized medicine and yearly we get the evidences from 40,000 trials [16]. However, we surely do not have enough evidences for all aspects of the practice of medicine,

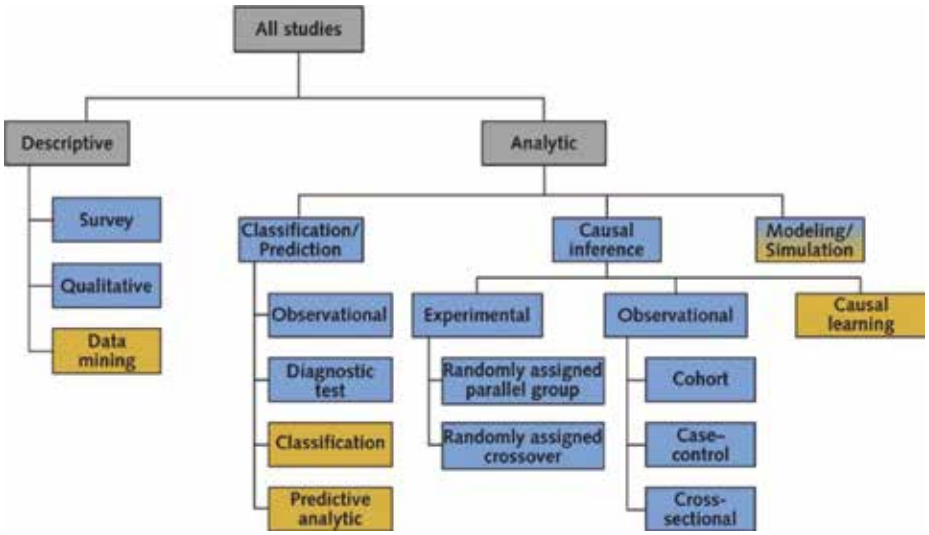

Figure 2. Big data (yellow) and randomized controlled trials (RCT = blue) [17].

the intuition is fading away and therefore, we have to look around for other solutions. Big data could be the problem solver [16]. The combination of RCTs and big data could be key to assure the quality and excellence in medicine in the future (**Figure 2**) [17].

4. Impact of modern machine learning in todays and future medicine practice

Machine learning (ML) technology already has a big impact on today's medical practice. From image classification in radiographs, over epidemic outbreak prediction to genome sequencing, computer algorithms become more and more prominent in modern medicine [18–21]. Projects of major companies like IBM Watson or Google's Deep Mind Health as well as numerous smaller privately and publicly funded research projects are pushing forward to close the gaps between medicine, mathematics, and computer science [22].

For narrow applications with clear regulations, ML algorithms already outperform human capabilities by far and even create new unseen strategies as recently shown by Google's Alpha Go Zero [22]. In this case, a self-trained algorithm mastered the game—more complex than chess—in less than 3 days. Although it has to be mentioned, that the reinforcement learning strategy used in this case might not be suitable in many medical applications, it is an indication of the potential of modern learning algorithms.

A closer look on the technology gives hope, that similar algorithms will not be limited to single and narrow applications, but rather can evolve in order to address the challenges of preserving the knowledge and intuition of experts and improve the quality of RCTs.

The reason for the great potential of machine learning lies in the nature of most of these new algorithms. Regardless whether Random Forests, Support Vector Machine or—most popular

today—Deep Neural Networks are applied, they generate their functionality by converting the information contained in thousands and even millions or billions of examples into a highly nonlinear mathematical model [23–25]. In some ways, this is similar to the human learning process and hence may be the appropriate tool for conserving human experience in such models.

With medical data becoming increasingly available in a digitalized form, not only by clinical trials but rather from every day medical practice, the databases grow in size and in depth [26]. This provides the possibility for algorithms not only to become more precise in their predictions but also to become more general in a sense that they can include a huge variety of factors in their decision process. Some aspects of a fading in intuition may so be replaceable by recommendation systems not based on thousands of hours of experience but of millions of decisions already made by experts in the past. Of course, so far, humans are still more efficient in their learning capabilities, but the pure scale of data contained in the algorithmic models may overcome this lack in efficiency. Nevertheless, recommendation systems driven by machine learning algorithms are more likely to complement a physician's intuition in the very near future, than to completely surrogate it.

In a similar manner, classical RCTs can benefit from big data. In semi-supervised learning, for example, datasets of known and unknown outcomes are considered. Here one task is to identify corner cases in the data in order to increase the model accuracy most efficiently [27]. Such methods could help to identify suitable candidates for clinical trials to make their results more robust. Another opportunity in this direction is the analysis of already trained models by feature extraction methods, which may generate promising hypotheses for further investigation in RCT. Finally it is to mention, that although RCTs form the scientific backbone of medicine, factors like publication bias and poor reproducibility rates show, that permanent monitoring of standard clinical practice is necessary [28, 29]. The ability of machine learning algorithms to constantly adapt to new situations, they seem to be predestined for such a controlling task.

5. Conclusion

The excellence in the practice of medicine was bonded to long working hours and a relative small knowledge base. Nowadays, the framework for the practice of medicine is protecting the rights of the individual regarding long working hours, however, in combination with the fast growing of the knowledge base, the practice of medicine is under pressure regarding quality and excellence. The big data approach could help find a solution; however, the digitalization of all data used in the practice of medicine before are warranted. In combination, IT will help to improve the quality of care.

Conflict of interest

There is no "conflict of interest" declaration necessary.

Author details

Mark Dominik Alscher[1]* and Nico Schmidt[2]

*Address all correspondence to: dominik.alscher@rbk.de

1 Robert-Bosch-Hospital, Stuttgart, Germany

2 Bosch - Healthcare Solution GmbH, Waiblingen, Germany

References

[1] Alscher MD. Medicine in the digital age: Let's take the opportunities! Deutsche Medizinische Wochenschrift. 2017;**142**(5):313

[2] Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Academic Medicine. 2004;**79**:S70-S81

[3] Galton F. Hereditary Genius: An Inquiry into Its Laws and Consequences. London: Originally published in 1869: Julian Friedman Publishers; 1979. p. 1869

[4] Simon HA, Chase WG. Skill in chess. American Scientist. 1973;**61**:394-403

[5] Mylopoulos M, Lohfeld L, Norman GR, Dhaliwal G, Eva KW. Renowned physicians' perceptions of expert diagnostic practice. Academic Medicine. 2012;**87**(10):1413-1417

[6] Croskerry P. A universal model of diagnostic reasoning. Academic Medicine. 2009;**84**: 1022-1028

[7] Flexner A. Medical education in the United States and Canada. Carnegie Foundation for the Advancement of Teaching. 1910; Bulletin No. 4

[8] Stern DT, Papadakis M. The developing physician—Becoming a professional. New England Journal of Medicine. 2006;**355**:1794-1799

[9] Greenberg A, Verbalis JG, Amin AN, Burst VR, Chiodo 3rd JA, Chiong JR, et al. Current treatment practice and outcomes. Report of the hyponatremia registry. Kidney International. 2015;**88**(1):167-177

[10] Bickel J, Brown AJ. Generation X: Implications for faculty recruitment and development in academic health centers. Academic Medicine. 2005;**80**:205-210

[11] Eckleberry-Hunt J, Tucciarone J. The challenges and opportunities of teaching "generation y". Journal of Graduate Medical Education. 2011;**3**(4):458-461

[12] Kühn S, Haas P. Elektronische Patientenakte: Von der Dokumentenakte zur feingranularen Akte. Dtsch Arztebl International. 2008;**105**(18):20

[13] Henderson ML, Dayhoff RE, Titton CP, Casertano A. Using IHE and HL7 conformance to specify consistent PACS interoperability for a large multi-center enterprise. Journal of Healthcare Information Management. 2006;**20**(3):47-53

[14] Heinze O, Birkle M, Koster L, Bergh B. Architecture of a consent management suite and integration into IHE-based regional health information networks. BMC Medical Informatics and Decision Making. 2011;**11**:58

[15] Windeler J, Antes G, Behrens J, Donner-Banzhoff N, Lelgemann M. Randomised controlled trials (RCTs). Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen. 2008;**102**(5):321-325

[16] Angus DC. Fusing randomized trials with big data: The key to self-learning health care systems? Journal of the American Medical Association. 2015;**314**(8):767-768

[17] Sim I. Two ways of knowing: Big data and evidence based medicine. Annals of Internal Medicine. 2016;**164**:262-263

[18] DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics. 2011;**43**(5):491-498

[19] Akgul CB, Rubin DL, Napel S, Beaulieu CF, Greenspan H, Acar B. Content-based image retrieval in radiology: Current status and future directions. Journal of Digital Imaging. 2011;**24**(2):208-222

[20] Kircher M, Stenzel U, Kelso J. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. Genome Biology. 2009;**10**(8):R83

[21] Penedo MG, Carreira MJ, Mosquera A, Cabello D. Computer-aided diagnosis: A neural-network-based approach to lung nodule detection. IEEE Transactions on Medical Imaging. 1998;**17**(6):872-880

[22] Gibney E. What Google's winning Go algorithm will do next. Nature. 2016;**531**(7594): 284-285

[23] Basu S, Kumbier K, Brown JB, Yu B. Iterative random forests to discover predictive and stable high-order interactions. Proceedings of the National Academy of Sciences of the United States of America. January 19, 2018;**2017**:11236. (published ahead of print)

[24] Goodfellow I, Yoshua B, Courville A. Deep Learning. Cambridge: MIT press; Springer-Verlag New York. 2016. http://www.springer.com/978-0-387-77241-7

[25] Steinwart I, Andreas C. Support Vector Machines. New York: Springer-Verlag, Springer Science & Business Media; 2008

[26] Bellazzi R, Blaz Z. Predictive data mining in clinical medicine: Current issues and guidelines. International Journal of Medical Informatics. 2008;**77**(2):81-97

[27] Cohn D, Rich C, McCallum A. Semi-supervised clustering with user feedback. Constrained Clustering: Advances in Algorithms, Theory, and Applications. 2003;**4**(1):17-32

[28] Mobley A, Linder SK, Braeuer R, Ellis LM, Zwelling L. A survey on data reproducibility in cancer research provides insights into our limited ability to translate findings from the laboratory to the clinic. PLoS One. 2013;**8**(5):e63221

[29] Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet. 1991;**337**(8746):867-872

Applications

Chapter 6

Use of Artificial Intelligence in Healthcare Delivery

Sandeep Reddy

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.74714

Abstract

In recent years, there has been an amplified focus on the use of artificial intelligence (AI) in various domains to resolve complex issues. Likewise, the adoption of artificial intelligence (AI) in healthcare is growing while radically changing the face of healthcare delivery. AI is being employed in a myriad of settings including hospitals, clinical laboratories, and research facilities. AI approaches employing machines to sense and comprehend data like humans has opened up previously unavailable or unrecognised opportunities for clinical practitioners and health service organisations. Some examples include utilising AI approaches to analyse unstructured data such as photos, videos, physician notes to enable clinical decision making; use of intelligence interfaces to enhance patient engagement and compliance with treatment; and predictive modelling to manage patient flow and hospital capacity/resource allocation. Yet, there is an incomplete understanding of AI and even confusion as to what it is? Also, it is not completely clear what the implications are in using AI generally and in particular for clinicians? This chapter aims to cover these topics and also introduce the reader to the concept of AI, the theories behind AI programming and the various applications of AI in the medical domain.

Keywords: artificial intelligence, healthcare delivery, medicine, machine learning, deep learning, intelligent agent and neural networks

1. Introduction

There has been an immense amount of discussion in recent years about the advent of artificial intelligence (AI) and the implication of its application in various domains. However, the concept of AI is not new and can be traced back to Ramon Llull's theory of a reasoning machine in 1300 CE and even Aristotle's syllogisms in 300 BC [1, 2]. However, it is only since the 1950s, clearer definitions and practical applications have been formulated [3, 4]. While

there was a lull in the development of AI in the 70s and 80s because of loss of interest and funding, there has been in the most recent period a dramatic revival in the research and development of AI programs. Countries like China have prioritised AI development by investing billions of dollars into AI industrial hubs [5]. Other nations and global corporations have also invested into AI programming and creation of innovative AI applications [6–8]. Building on this trend, institutions are now increasingly paying attention to application of AI in healthcare. AI is being used to improve the efficiency in delivery of healthcare and address previously intractable health problems [1, 9, 10]. The hundreds of AI-based healthcare applications being introduced into the market in recent years is a testimonial to this focus. Commentators have discussed how application of AI in healthcare is at the early stages and there is yet more to come [1, 4, 6]. However, is AI just hype and are entities investing into a bubble? To get an answer, we first need to understand what AI is and its approaches and tools. This chapter covers these issues and how they specifically apply to healthcare and what is next for the use of AI in healthcare?

2. Development and application of AI

2.1. Definition

So, what is AI? Because of the complexity involved in developing synthetic intelligence that is comparable to human intelligence, there are varying interpretations of what AI is and what goes into developing AI. Some authors even frown upon the term 'AI' and prefer the term 'Computational Intelligence' [11]. However, if we consider what is the objective of AI and what resources go into achieving the objective, an acceptable definition encompassing these components can be fashioned. The end objective of AI is to create systems that think and act rationally like humans [2, 4, 12]. These systems can also be termed as 'intelligent agents' [2, 4]. If the goal of the system is to demonstrate intelligence and developing these systems requires computer programming, a formal definition of AI would read as '*a field of science concerned with the computational understanding of what is commonly called intelligent behaviour, and with the creation of intelligent agents that exhibit such behaviour*' [13]. Simpler definitions describe AI as 'machines assuming human like capabilities', 'extension of human intelligence through computers' and 'making computers do things which currently humans do' but a more accurate description would be 'the science of making intelligent machines' [1, 2, 4, 14].

2.2. Intelligent agent

AI theory can be best understood through the *intelligent agent* concept [11]. An intelligent agent incorporates the skills required to pass the Turing Test, which assesses whether a machine can think like a human? [2, 3]. So an intelligent agent should be skilled in perception, practical reasoning and have an ability to take action to achieve its goals. The agent utilises the environment, it operates within, to both receive input and take action (**Figure 1**). Some key inputs that feed into an agent and potentially, which it can draw itself are current observations about the environment, prior knowledge about the environment, past experiences

http://dx.doi.org/10.5772/intechopen.74714

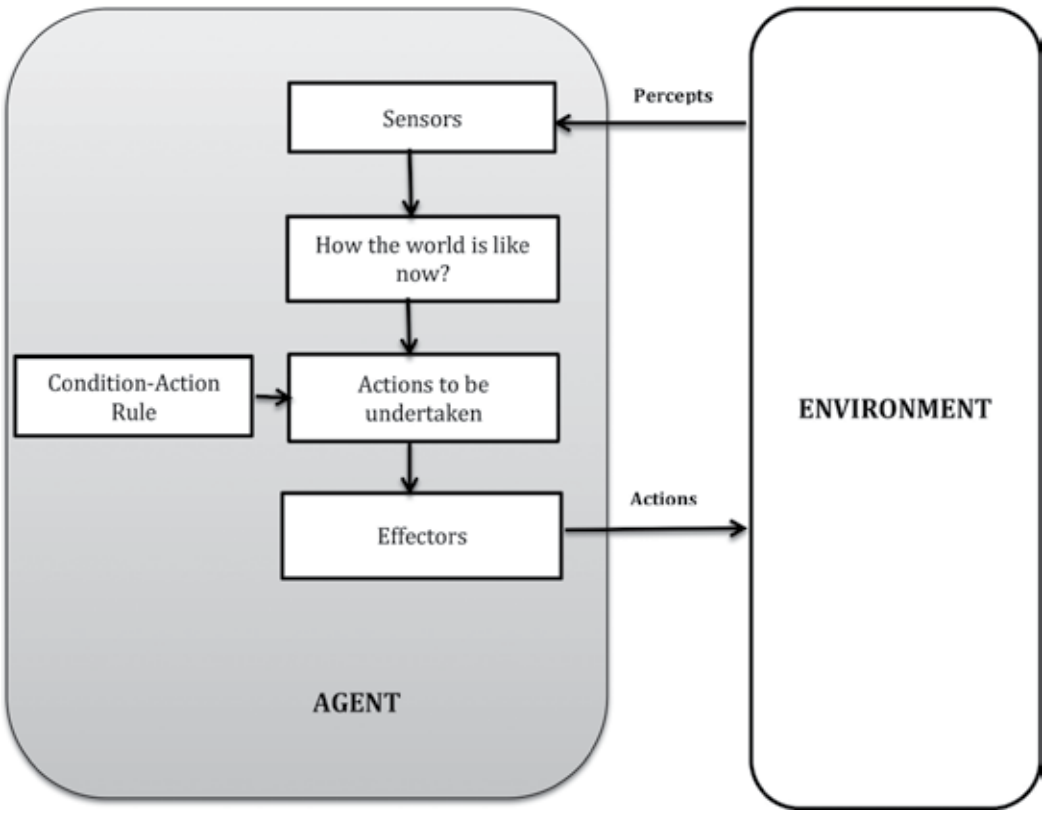

Figure 1. Concept of an intelligent agent.

that it can learn from and the objectives it needs to achieve. The agent perceives the environment through sensors and acts on the environment through effectors. When an intelligent agent is comprised of a computational core with physical actuators and sensors, it is termed a 'robot' [11]. When an agent is a program acting in a pure computational environment, it is an 'infobot' and when an advice providing program is coupled with a human expert, it is a 'decision support system'.

2.3. What makes up AI?

In the past, researchers aimed for AI to replicate human intelligence [2]. This approach is called 'Classical AI'. However, this was a limiting approach as it assumed human intelligence is the only form of intelligence. This approach also assumes human intelligence is the most intelligence can be. Intelligence mainly comprises of learning and reasoning [3, 13]. Constructing intelligence does not have to be defined by the limitations human intelligence poses. An apt analogy to discuss here is flight. While bird flight may be a source of inspiration for constructing aeroplanes, the aeroplane structure is not replicating the anatomic structure of a bird. So in constructing AI, it is more important to incorporate the vital characteristics of intelligence than merely replicate human intelligence.

Learning is an essential characteristic of intelligence [2, 4, 11]. Learning involves acquiring new knowledge, developing new skills through instruction or practise and knowledge representation and experimentation. If AI comprises learning, it has to demonstrate all the aforementioned features. A very common process through which AI systems achieve learning objectives is by *Machine Learning*. Machine learning is the modelling of different aspects of the learning process by computers [15]. Key goals of machine learning are for algorithms[1] to self-learn and improve through experience. The algorithms for machine learning typically fall into two categories: supervised and unsupervised categories [16]. Supervised learning involves an algorithm working with labelled training data. Categorisation of data and programming of the relationship between input and output data occurs in supervised learning. On the other hand, unsupervised learning allows the algorithm to identify a hidden pattern in a stack of data. Here the algorithm is run to check what patterns can be identified in the data and what outcomes may occur?

Reasoning and knowledge representation are the other aspects of AI [11]. In AI, reasoning involves manipulation of data to produce actions. Unlike traditional programming, the emphasis in AI is on what is to be computed rather than how it is to be computed? Structuring of this computation happens through design-time reasoning, offline computation and online computation. Earlier forms of AI involved algorithms based on the step-by-step reasoning model used to address predicated problems [2]. However, these models were not useful for uncertain situations or when there was incomplete information. AI reasoning models have now evolved to respond to these situations by drawing upon concepts from probability and economic theories. To resolve problems-certain or uncertain, AI systems require widespread knowledge about the relevant environment and then be able to represent this knowledge in a computable form [11]. For this to occur, AI uses a *Representation and Reasoning System* (RRS). An RRS is comprised of a programming language to communicate with a computer, a method to allocate meaning to the language and after input a process to figure out the answers. Knowledge is represented in different forms, but the most widely used method is Frames [2]. Frames are files in the computer where information is stored in slots. To enable AI knowledge representation and reasoning, programming languages and computational resources are two important properties. Different programming languages are used in AI, but the most popular are low-level programming languages such as Lisp, Python, C++ and Fortran. In the past, stand-alone computers and their limited processing power had restricted the advancement of AI. In recent years, AI reasoning and knowledge representation has immensely benefited from the rapid technological advances in computing power and wireless technology. These advances have helped in the deployment of sophisticated algorithms designed to resolve problems that could not have been addressed by AI applications in the past.

2.4. AI tools

AI systems employ several tools to automate problem-solving tasks. These tools are based on AI principles, some of which were discussed in the previous sections. The tools are used to

[1]In computer science, an algorithm is an explicit description of how to solve a class of problems? [2–4].

create AI applications to resolves issues across various disciplines and industries. Some commonly utilised tools are discussed in this section.

Search in AI system mirrors real-life problem solving but draws upon computing power to resolve the problems [17]. Search problems are classified based on the amount of information that is available to the search process. This information may relate to the whole of the problem area or a specific component of the problem. AI through an independent search planning process analyses multiple options and identifies an optimal solution. AI adopts a faster and better process to search and optimisation than conventional techniques [17, 18]. The search process that separates AI from conventional techniques is its process remembers past results, learns and refines its performance in relation to past searches, plans its path forward and answers search queries akin to human intelligence. One such example of AI search and optimisation tool is *Evolutionary Computation*. Evolutionary Computation is the umbrella term for algorithms based on natural evolutionary processes that incorporate mechanisms of natural selection and survival of the fittest principle [1, 10]. Foremost of the evolutionary computation algorithms are the *Genetic Algorithms*. Genetic Algorithms are a category of stochastic search and optimisation algorithms based on Darwin's natural biological evolution. These algorithms use a population-based search process to create random solutions for the problem at hand. These solutions are termed chromosomes. The chromosomes are comprised of random values derived from various control values. The variations in the values are utilised for the search process. The population of chromosomes is then assessed for an objective function. This population of solutions then evolves from one generation to another to arrive at an acceptable solution. The ideal solutions are retained and the mediocre ones disposed of. Through a process of repetition, improvements and generation of new solutions would occur.

In their quest to replicate biological intelligence, AI researchers inspired by the biological nervous system have developed *Artificial Neural Networks (ANNs)* [1, 19]. Artificial Neural Networks attempt to simulate nerve cell (neurons) networks of the brain. This approach of copying biological neuronal networks to function independently differs from conventional computing process that primarily seeks to support human brain computation. A very simple base algorithm structure (see **Figure 2**) lies behind the artificial neural networks but it can be adapted to a range of problems. The artificial neurons, which are computer processors, are interconnected with each other and are capable of performing parallel computations for data processing and knowledge representation [19]. These neural networks are capable of learning from historical examples, examining non-linear data, and managing imprecise information. ANNs are categorised into two main categories: *Feedforward Neural Networks* and *Recurrent Neural Networks*. In feedforward network the signal passes in only one direction and in recurrent neural networks, feedback and short-term memories of previous inputs are enabled. In both categories, application of deep learning, which is a class of machine learning that uses a cascade of multiple layers of non-linear processing units, enhances the problem-solving capabilities of the neural networks. So we have deep feedforward and deep recurrent neural networks increasingly being used to resolve real-world problems through language modelling, analysis of unstructured data and strategy formulation.

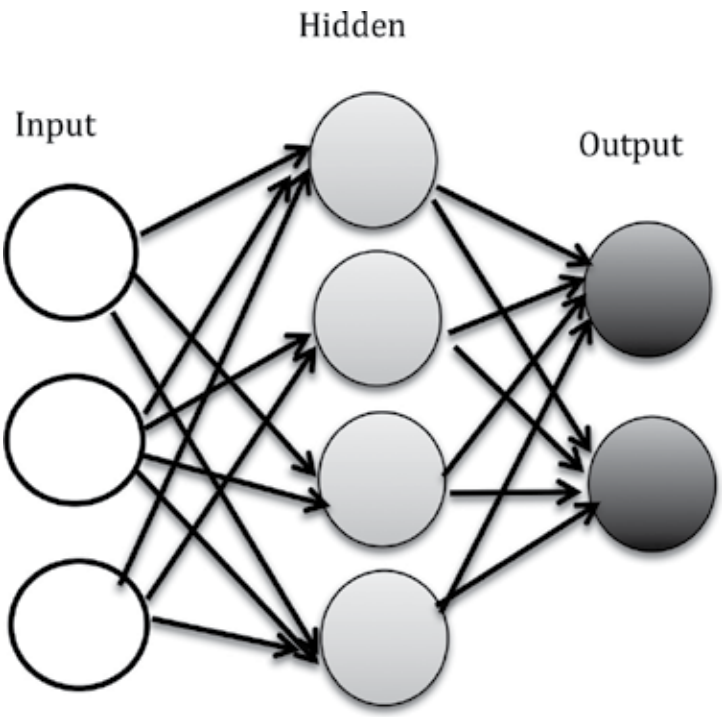

Figure 2. Schematic representation of an artificial neural network.

Logic is important to reasoning, which in turn is a key component of intelligence. Classical logic is based on the assumption that only two truth-values (false and true) exist [2]. This assumption is called *bivalence*. On the other hand, *Fuzzy Logic* reflects real world phenomenon, where everything is a matter of degree [1, 2, 10]. A fuzzy logic can be viewed as a fuzzy extension of a multi-valued logic i.e. instead of recognising everything is black and white, it recognises there are shades of grey. Fuzzy logic uses continuous set membership from 0 to 1 in opposition to Boolean logic, which relies on sharp distinctions such as 0 for false and 1 for true. Fuzzy applications utilise a structure of series of 'if-then' rules for modelling. This approach by fuzzy logic permits ambiguity and can be used in AI systems for indeterminate reasoning.

Another important AI technique is *Natural Language Processing*. Natural language processing (NLP) is concerned with the use of software programming to understand and manipulate natural language text or speech for practical purposes [20]. With NLP, the process of language analysis is decomposed into various stages mirroring theoretical linguistic distinctions as outlined by syntax, semantics and pragmatics [4, 20]. NLP enables machines to read and understand human language. NLP can also be utilised to gather and analyse unstructured data such as free text. In recent years, progress in NLP specifically in the field of syntax has led to development of effective grammar characterisation and chart parsing. Development of numerous conceptual tools has led to formation of systems and interface subsystems to use

for experiments. Part-of-speech identification and word sense disambiguation have become standard processes in NLP. Other current applications of NLP include information retrieval, machine translation and text mining.

Use of *Hybrid Artificial Intelligent Systems* (HAIS), which are a combination of AI techniques, is becoming popular because of its capabilities to address real world complex problems that individual AI techniques cannot address [21]. By combining different AI learning and adaptation techniques, HAIS overcomes the limitations associated with a particular technique. HAIS may involve a combination of agents and multi-agent systems, fuzzy systems, artificial neural networks, optimisation models and so forth. By combining symbolic and sub-symbolic techniques, complex issues involving indistinctness, ambiguity and vagueness can be resolved by HAIS. The synergy in HAIS also allows it to adjust to common sense, mine knowledge from raw data, use human like reasoning, and learn to adapt to a changing environment.

3. AI in healthcare

AI lends itself to healthcare delivery very well. In fact, in the recent years there has been an exponential increase in the use of AI in clinical environments [1, 6, 21–24]. With modern Medicine facing a significant challenge of acquiring, analysing and applying structured and unstructured data to treat or manage diseases, AI systems with their data-mining and pattern recognition capabilities come in handy. Medical AI is mainly concerned with the development of AI programs that help with the prediction, diagnosis and treatment or management of diseases. In contrast to non-AI medical software application, which relies on pure statistical analysis and probabilistic approaches, medical AI applications utilise symbolic models of diseases and analyse their relationship to patient signs and symptoms [1, 25–27]. For example, diagnostic AI applications gather and synthesise clinical data and compare information with predefined categories such as diseases to help with diagnosis and treatment. Medical AI applications have not just been used to support diagnosis but also treatment protocol development, drug development and patient monitoring too [1].

3.1. History of use of AI in healthcare

Discussion of the use of AI in medicine coincides with the advent AI in the modern era. This is not surprising as AI systems initially intend to replicate the functioning of the human brain [2]. In 1970, William B Schwartz, a physician interested in the use of computing science in medicine, published an influential paper in the *New England Journal of Medicine* titled '*Medicine and the computer: the promise and problems of change*' [28]. In the paper he argued '*Computing science will probably exert its major effects by augmenting and, in some cases, largely replacing the intellectual functions of the physician'*. By the 1970s there was a realisation that conventional computing techniques were unsuitable for solving complex medical phenomenon [2, 4]. A more sophisticated computational model that simulated human cognitive processes, that is AI models, was required for clinical problem solving. Early efforts to apply AI in medicine consisted of setting up rules-based systems to help with medical reasoning. However, serious

clinical problems are too complex to lend them to simple rules-based problem solving techniques. Problem solving in medicine then progressed to construction of computer programs based on models of diseases. It was not just with the field of general medicine, that AI was being explored to assist with problem solving. In 1976, the Scottish surgeon Gunn used computational analysis to diagnose acute abdominal pain [1]. This was achieved through clinical audits of structured case notes through computers, whereby diagnosis through this route proved to be about 10% more accurate than the conventional route. By the 1980s, AI research communities were well established across the world but especially in learning centres in the US [1, 2, 4, 13]. This development helped in expansion of the use of novel and innovative AI approaches to medical diagnoses. Much of this push was because medicine was an ideal testing ground for these AI applications. A significant number of AI applications in medicine at this stage were based on the *expert system* methodology [1, 25, 29–31]. By the end of the 1990s, research in medical AI had started to use new techniques like machine learning and artificial neural networks to aid clinical decision-making. The next section explores current application of AI in various aspects of healthcare.

3.2. Application of AI techniques in healthcare

The wide acceptance of AI in healthcare relates to the complexities of modern medicine, which involves acquisition and analysis of the copious amount of information and the limitation of clinicians to address these needs with just human intelligence. Medical AI applications with their advanced computing ability are overcoming this limitation and are using several techniques to assist clinicians in medical care.

AI is being used for all the three classical medical tasks: diagnosis, prognosis and therapy but mostly in the area of medical diagnosis [9, 32]. Generally, the medical diagnosis cycle (**Figure 3**) involves observation and examination of the patient, collection of patient data, interpretation of the data using the clinician's knowledge and experience and then formulation of a diagnosis and a therapeutic plan by the physician. If we can compare the medical diagnostic cycle (**Figure 3**) to the concept of an intelligent agent system, the physician is the intelligent agent, the patient data is the input and the diagnosis is the output. There are several methods, through which AI systems can replicate this diagnostic cycle and assist clinicians with medical diagnosis. One such approach is the use of *Expert Systems*. Expert systems are based on rules clearly outlining the steps involved in progressing from inputs to outputs [2]. The progression occurs through the construction of a number of IF-THEN type rules. These rules are constructed with the help of subject experts like clinicians who have interest and experience in the particular domain. The success of the expert system relies on the explicit representation of the knowledge area in the form of rules. The core of the expert system is the inference engine, which transforms the inputs into actionable outputs.

Commonly, the application of the expert system approach in medical software programming is seen in *Clinical Decision Support Systems* (CDSS). Simply put, CDSS are software programs that enable clinicians to make clinician decisions [33, 34]. CDSS provides customised assessment or advice based on analysis of patient data sets. An early version of CDSS was the MYCIN program developed in the 1970s. MYCIN was a CDSS focusing on the management

http://dx.doi.org/10.5772/intechopen.74714

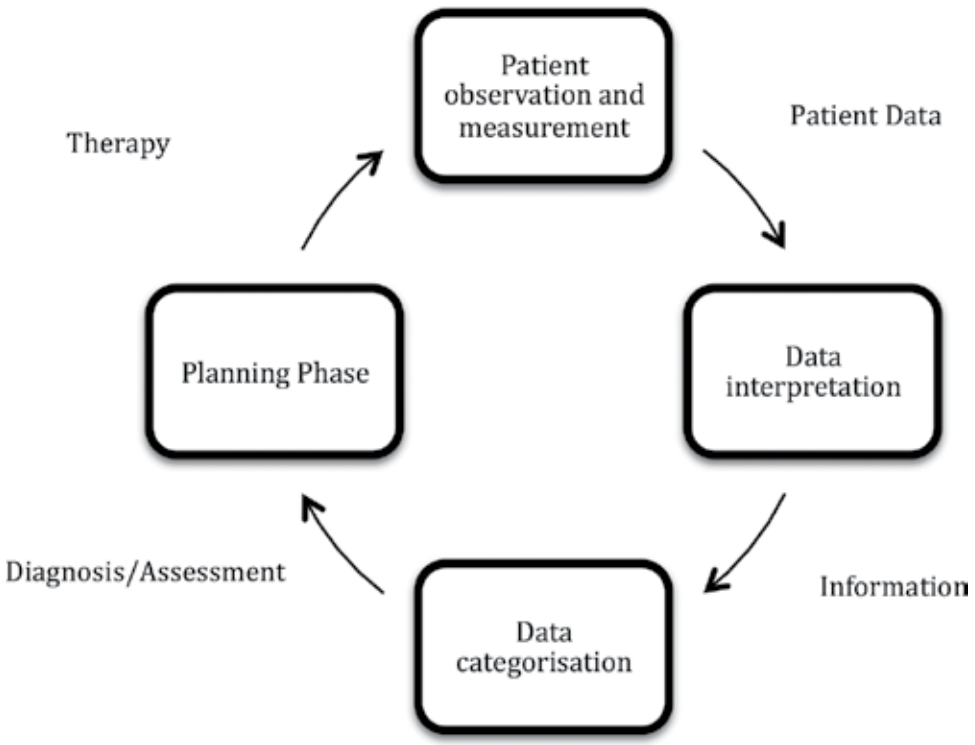

Figure 3. Medical diagnostic-therapeutic cycle.

of infectious disease patients. Infectious disease knowledge was represented in the form of production rules, which are conditional statements as to how observations can be inferred appropriately. However, MYCIN had less emphasis on diagnosis and more on the management of patients with infectious diseases. In a later evaluation of the MYCIN system, it was found it compared favourably with the advice provided by infectious disease experts. MYCIN paved the way for the development of knowledge-based systems and the commercialisation of rule-base approaches in medicine and other fields. Another CDSS that was initially developed around the same time period as MYCIN but continues to be used is the QMR system [35]. The QMR system utilises a customised algorithm modelled on the clinical reasoning of one single University of Pittsburgh internist. Hence the system was initially called INTERNIST-I. By considering historical and physical findings, QMR system generates the differential diagnosis. Utilising a large database that categorises disease findings into 'evoking strengths', 'importance' and 'frequencies' domains, the system generates the differential diagnosis. Heuristic rules drove the system to produce a list of ranked diagnoses founded on disease knowledge domains in-built into the system. Where the system was unable to make a determined diagnosis it probed the user with further questions or provided advice about further tests until a determination of the condition was made. While MYCIN and QMR systems offered diagnostic support, other forms of CDSS can provide alerts and reminders and advice about patient treatment and management. These systems operate by

creating predictive models and multi-dimensional patient view through aggregation of data from multiple sources including knowledge and patient information databases. As treatment and management of diseases have evolved, CDSS architecture is now utilising multi-agent systems [26]. Each of the multiple agents performs distinct tasks and operations in various capacities or different locations but transmit data to a central repository so aggregated data can be used for knowledge discovery.

Unlike experts systems where a serial or sequential data processing approach is utilised, ANN processing utilises a parallel form of data processing analogous to the brain [19]. In ANNs, the processing elements, otherwise called as neurons, process data simultaneously while communicating with each other. The processing elements are arranged in layers and the layers, in turn, are connected to each other. The links between the processing elements are associated with a numerical weight. The memory and adaptation of ANNs are adjusted by changing the weights, which leads to the amplification of the effects of afferent connection to each processing element. As a result of this architecture, ANNs can be trained to learn from experience, analyse non-linear data and manage inexact information. These abilities have led to ANN techniques being one of the most popularly utilised AI techniques in medicine [1]. ANNs in addition to medical diagnosis have been used for radiology and histopathology analysis. In radiology, gamma camera, CT, ultrasound and MRI all create digital images, which can be manipulated by ANNs and used as inputs. The digitised inputs are then transmitted through the hidden and output layers to produce desired outputs (see **Figure 2**). Using the Backpropagation approach, a learning algorithm, ANNs have successfully identified orthopaedic trauma from radiographs [36]. When ANNs and radiologists interpret the same radiological images separately, research has identified good diagnostic agreement [1, 36]. ANNs have also been used for analysis of cytological and histological specimens too [1, 25]. For example, ANNs has been used to screen abnormal cells from slide images for haematology and cervical cytology. Further, ANNs have also been used to interpret ECGs and EEGs through waveform analysis. For this to occur, a multi-layered neural network is trained with waveform data from both people with the disease and without [1]. Evaluation of the waveform interpretations by ANNs has identified excellent pattern approximation and classification abilities and comparable in interpretation to clinicians.

Data Mining acts as the foundation for machine learning. Data mining is the process for identifying previously unknown patterns and trends in large databases and then utilising the same to create predictive models [37, 38]. Data mining involves multiple iterative steps (**Figure 4**) that includes retrieval of data sets from data warehouses or operational databases, cleaning of data to remove discrepancies, analysis of data sets to identify patterns that represent relationships amongst the data, validation of the patterns with new data sets and culminating in knowledge extraction [39]. Use of data mining has become hugely popular in healthcare largely because of the generation of data too voluminous and complex to be processed by conventional computational techniques. The potential application of data mining in healthcare can be huge but practically data mining has been used in evaluating the effectiveness of medical treatments, analyse epidemiological data to identify disease outbreaks and act as an early warning system, analyse hospital records to identify acute medical conditions and help with interventions, quality assessment of medical interventions and predicting survival

http://dx.doi.org/10.5772/intechopen.74714

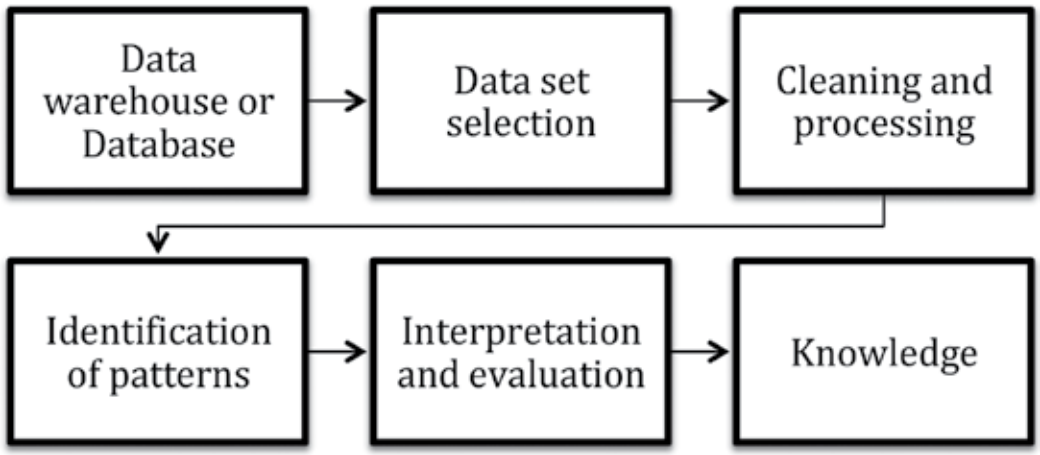

Figure 4. Data mining process. Adapted from Huang et al. [39].

time for chronic disease and cancer patients [8, 38–40]. Data mining medical data faces two main issues: heterogeneity of data sometimes with incomplete recording or filing of data and complexity of the requested outputs [27]. Fuzzy logic, which we discussed in an earlier section, with its proficiency to represent assorted data, strength in adapting to change in the user environment and its distinctive expressiveness can support data mining in addressing these issues. Thus data mining utilising fuzzy logic has been used for a range of situations in healthcare including prediction of the prognosis of cancer and assessing the satisfaction of clinicians for patient information management systems.

There are an estimated 5 billion mobile phone subscriptions in the world [41]. Many mobile phones now have memories and processing power equivalent to the capacity of mini-computers [42]. So it is natural to see mobile communication devices being harnessed to deliver healthcare. The use of wireless communication devices to support delivery of healthcare is called *Mobile Health* or in a popular terminology: *mHealth* [41]. Mobile health applications are being used in many areas of healthcare delivery including education and awareness, point-of-care support and diagnostics, patient monitoring, disease surveillance, emergency medical response and patient information management [41, 43–46]. The rapid development in mHealth has coincided with the increase in AI research and development of AI techniques. Consequently, there has been an increased application of AI techniques in mHealth. The move has worked well as characteristics of an intelligent agent system lend themselves to the objectives of mHealth. The intelligent agent perceives the environment and autonomously acts upon it. In case of multi-agent systems, the agents can communicate between themselves, dynamically manage data and resource and handle the complexity of solutions through decomposition, modelling and reorganisation of relationships. These abilities mean agent-based mobile applications can be used for remote monitoring of patients especially elderly and chronic disease patients, support clinical decision-making and provide remote training for health workers. The application of AI has not been restricted to mobile communication devices but has been extended to other smart devices. When these smart devices are

connected to each other to create a cyber-physical smart pervasive network, it is termed as the *Internet of Things* (IoT) [47, 48]. IoT is being used across for many purposes including prediction of natural disasters, water scarcity monitoring and intelligent transport systems but in health care, the concept is being used to design smart homes to assist senior citizens to accomplish their daily living activities while preserving their privacy and to remotely monitor their health conditions and medicine intake [48]. An IoT powered by AI and set up to address the healthcare need of senior and incapacitated patients is called as *Ambient Assisted Living* (AAL) [49]. As the main aim of AAL is to extend the independent living of elderly individuals in their homes, automation, security, control and communication are key aspects of AAL modular architecture. The system also includes sensors, actuators and cameras to collect different types of data about the individual and home. The constituent system sets up a smart home environment where activities and the health condition of the resident are not only tracked but also predicted [50].

In addition to the examples discussed above, AI techniques have been successfully used in other areas of medicine. Genetic algorithm techniques have been used to predict outcomes in acutely ill and cancer patients, to analyse mammograms and MRI images and fuzzy logic techniques have been used in diagnosing various cancers, characterise ultrasound and CT scan images and predict survival in cancer patients and administer medication and anaesthetics [1, 6].

Of all the AI applications that have been developed over the past many decades, IBM's Watson is one of the well-recognised applications. IBM Watson is a cognitive computing technology that groups together the competencies of reading, reasoning and learning to reply to questions or investigate original connections [40]. IBM Watson aggregates huge volumes of structured and unstructured data from multiple sources into a single repository called Watson corpus. IBM incorporates machine learning and NLP techniques to process and analyse data to undertake problem solving. The technology of IBM Watson has been extended to the medical domain to assist medical scientists and clinicians in improving patient care [31, 51–53]. Some of the published examples of the use of IBM Watson in health care include automated problem list generation from electronic medical records, drug target identification and drug repurposing, interpretation of genetic testing results, oncological decision making support, and to support the roll-out of government healthcare programs.

3.3. Future trends and application of AI in healthcare

As more AI research is undertaken and AI systems become more trained and consequently intelligent, it is foreseeable that these agents replace some of, if not all, the human elements of clinical care [6]. While leaving the communication of serious matters and final decision making to human clinicians, AI systems can take responsibility for routine and less risky diagnostic and treatment processes. The intention here is not to replace human clinicians but enable a streamlined high-quality healthcare delivery process.

Of all the promising medical AI novelties that are being explored, robotics driven by AI will have an important role in the medical automation process. Robots embody AI and give it a

form, while AI algorithms/programming provide intelligence to the robots [2]. Robotic assistants have already been employed to conduct surgeries, deliver medication and monitor hospital patients but the most promising area for their use is in elderly care [31]. Mobile robotic assistants are already being used to assist the elderly people in their day-to-day activities either in their home or in aged care settings [51]. The robotic assistants mainly undertake tasks that remind them of their routine activities including medication intake or guidance in their environments. With advances in AI and robotics, the employment of robotic assistants in elderly care is only bound to grow.

While the conventional thinking is that robots act as a vessel for a silicon-based artificial brain, there is emergence of a school of thought that imagines the use of biological brains in robots [2]. With advances in science now allowing the culture of biological neurons, the potential use of a biological brain in a robotic frame through which it can sense the world and move around is not inconceivable. This *Cyborg* model presents a true blurring of the boundaries between human and artificial intelligence and the imaginable development of a hybrid human-artificial intelligence health worker that can revolutionise healthcare delivery.

3.4. Challenges

While the application of AI in delivery of healthcare has very promising potential, challenges-both technical and ethical exist. AI research is largely led and driven by computer scientists without medical training and it has been commented that this has led to a very technologically focused and problem oriented approach in the application of AI in healthcare delivery [24]. Contemporary healthcare delivery models are very dependent on human reasoning, patient-clinician communication and establishing professional relationships with patients to ensure compliance. These aspects are something AI cannot replace easily. Use of robotic assistants in healthcare has raised issues about the mechanisation of care in vulnerable situations where human interaction and intervention is probably more appealing [6]. There is also the reluctance of clinicians in adopting AI technologies that they envisage will eventually replace them. Yet there is no qualm in them using technologies that automate and speed up laboratory diagnostic process [1]. This has led to some suggesting a model of co-habitation [6]. This is a model that accommodates both the AI and human elements in healthcare delivery and anticipates the inevitable automatisation of significant components of medical processes while preserving the human aspects of clinical care like communication, procedures and decision-making.

4. Conclusion

Healthcare delivery has over years become complex and challenging. A large part of the complexity in delivering healthcare is because of the voluminous data that is generated in the process of healthcare, which has to be interpreted in an intelligent fashion. AI systems with their problem solving approach can address this need. Their intelligent architecture, which incorporates learning and reasoning and ability to act autonomously without requiring constant human

attention, is alluring. Thus the medical domain has provided a fertile ground for AI researchers to test their techniques and in many instances; AI applications have successfully solved problems with outcomes comparable to that of human clinicians. As healthcare delivery becomes more expensive, stakeholders will increasingly look to solutions that can replace the expensive elements in patient care and AI solutions will be sought after in these situations. However, cold technology cannot totally replace the human elements in patient care and a model that incorporates both technological innovations and human care has to be investigated.

Notice

The chapter was submitted to a double blind review and it is in line with COPE Ethical Guidelines.

Author details

Sandeep Reddy

Address all correspondence to: sandeep.reddy@deakin.edu.au

School of Medicine, Faculty of Health, Deakin University, Australia

References

[1] Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. Annals of the Royal College of Surgeons of England. 2004;**86**(5):334-338

[2] Warwick K. Artificial Intelligence: The Basics. Abingdon: Routledge; 2012

[3] Simmons AB, Chappell SG. Artificial intelligence—Definition and practice. IEEE Journal of Oceanic Engineering. 1988;**13**(2):14-42

[4] Kok JN, Boers EJW, Kosters WA, Van Der Putten P, Poel M. Artificial Intelligence: Definition, Trends, Techniques, and Cases. Oxford, UK: Encyclopedia of Life Support Systems; 2013

[5] Cyranoski D. 'China enters the battle for AI talent'. Nature [Online]. 2018. Available from: https://www.nature.com/articles/d41586-018-00604-6 [Accessed: Jan 12, 2018]

[6] Diprose W, Buist N. Artificial intelligence in medicine: Humans need not apply? The New Zealand Medical Journal. 2016;**129**(1434):73-76

[7] Baker S. Final Jeopardy: The Story of Watson, the Computer That Will Transform Our World. New York: Houghton Mifflin Harcourt Publishing Company, Mariner Books; 2012

[8] Chen HCH, Zeng D. AI for global disease surveillance. IEEE Intelligent Systems. 2009; **24**(6):66-82

[9] Szolovitz P. Artificial intelligence in medical diagnosis. Annals of Internal Medicine. 1988; **108**(1):80

[10] Gambhir S, Malik SK, Kumar Y. Role of soft computing approaches in healthcare domain: A mini review. Journal of Medical Systems. 2016;**40**(12):2-20

[11] Poole DL, Mackworth A, Goebel RG. Computational intelligence and knowledge. In: Computational Intelligence: A Logical Approach. New York: Oxford University Press; no. Ci. 1998. pp. 1-22

[12] Poole DL, Mackworth AK. Artificial Intelligence: Foundations of Computational Agents. 2nd ed. Cambridge, UK: Cambridge University Press; 2017

[13] Shapiro SC. Encyclopedia of Artificial Intelligence. 2nd ed. New York: Wiley-Interscience; 1992

[14] Szolovits P, Stephen PG. Medicine Flowchart—How Can We Emulate Diagnosis? Categorical and Probabilistic Reasoning in Medical Diagnosis; Artificial Intelligence. 1978;**11**:115-144

[15] Bell J. Machine learning. In: Machine Learning: Hands-on for Developers and Technical Professionals. Indianapolis, USA: John Wiley & Sons; 2015

[16] Michalski RS, Carbonell JG, Mitchell TM. An overview of machine learning. In: Michalski RS, Carbonell JG, Mitchell TM, editors. Machine Learning: An Artificial Intelligence Approach. Palo Alto: Springer; 1983. p. 23

[17] Chandel A, Sood M. Searching and optimization techniques in artificial intelligence: A comparative study and complexity analysis. International Journal of Advanced Research in Computer Engineering & Technology. 2014;**3**(3):866-871

[18] Badar A, Umre BS, Junghare AS. Study of artificial intelligence optimization techniques applied to active power loss minimization. IOSR Journal of Electrical and Electronics Engineering. 2014;**2014**(1989):39-45

[19] Priddy KL, Keller PE. Artificial Neural Networks: An Introduction. Bellingham: SPIE Press; 2005

[20] Spyns P. Natural language processing in medicine: An overview. Methods of Information in Medicine. 1996;**35**(4-5):285-301

[21] Abraham A. Hybrid artificial intelligent systems. In: Corchado E, Corchado J, Abraham A, editors. Innovations in Hybrid Intelligent Systems—Advances in Soft Computing. Berlin: Springer Berlin Heidelberg; 2007

[22] Kim E-Y. Patient will see you now: The future of medicine is in your hands. Healthcare Informatics Research. 2015;**21**(4):321-323

[23] Mishra S, Takke A, Auti S, Suryavanshi S, Oza M. Role of artificial intelligence in health care. BioChemistry: An Indian Journal. 2017;**11**(5):1-14

[24] Coiera EW. Artificial intelligence in medicine: The challenges ahead. Journal of the American Medical Informatics Association. 1996;**3**(6):363-366

[25] Scott R. Artificial intelligence: its use in medical diagnosis. Journal of Nuclear Medicine. 1993;**34**(3):510-514

[26] Wimmer H, Yoon VY, Sugumaran V. A multi-agent system to support evidence based medicine and clinical decision making via data sharing and data privacy. Decision Support Systems. 2016;**88**:51-66

[27] Koh HC, Tan G. Data mining applications in healthcare. Journal of Healthcare Information Management. 2005;**19**(2):64-72

[28] Schwartz WB. Medicine and the computer: The promise and problems of change. The New England Journal of Medicine. 1970;**283**(23):1257-1264

[29] Sim I, Gorman P, Greenes RA, Haynes RB, Kaplan B, Lehmann H, Tang PC. Clinical decision support systems for the practice of evidence-based medicine. Journal of the American Medical Informatics Association. 2001;**8**(6):527-534

[30] Khanna S, Sattar A, Hansen D. Advances in artificial intelligence research in health. The Australasian Medical Journal. 2012;**5**(9):475-477

[31] Topol E. The Patient Will See You Now: The Future of Medicine Is in Your Hands. New York: Basic Books; 2015

[32] Farrugia A, Al-Jumeily D, Al-Jumaily M, Hussain A, Lamb D. Medical diagnosis: Are artificial intelligence systems able to diagnose the underlying causes of specific headaches? In: Proceedings of 2013 6th International Conference on Developments in eSystems Engineering (DeSE); 2013. pp. 376-382

[33] Littenberg B, MacLean C, Gagnon M. Clinical decision support system. US Pat. App. 11/640,103; 2006. pp. 1-10

[34] Pusic M, Ansermino JM. Clinical decision support systems. The British Columbia Medical Journal. 2004;**46**:236-239

[35] Miller RA, McNeil MA, Challinor SM, Masarie FE Jr, Myers JD. INTERNIST-1: An experimental computer-based diagnostic consultant for general internal medicine. The New England Journal of Medicine. 1982;**307**:468-476

[36] Olczak J, Fahlberg N, Maki A, Razavian AS, Jilert A, Stark A, Sköldenberg O, Gordon M. Artificial intelligence for analyzing orthopedic trauma radiographs: Deep learning algorithms—Are they on par with humans for diagnosing fractures? Acta Orthopaedica. 2017;**88**(6):581-586

[37] Milovic B, Milovic M. Prediction and decision making in health care using data mining. International Journal of Public Health Science. 2012;**1**(2):69-76

[38] Zhang Y, Guo SL, Han LN, Li TL. Application and exploration of big data mining in clinical medicine. Chinese Medical Journal. 2016;**129**(6):731-738

[39] Huang MJ, Chen MY, Lee SC. Integrating data mining with case-based reasoning for chronic diseases prognosis and diagnosis. Expert Systems with Applications. 2007;**32**(3):856-867

[40] Kelly JE III, Hamm S. Smart Machines: IBM's Watson and the Era of Cognitive Computing. New York: Columbia University Press; 2013

[41] Barton AJ. The regulation of mobile health applications. BMC Medicine. 2012;**10**:2-5

[42] Dagon D, Martin T, Starner T. Mobile phones as computing devices: The viruses are coming! IEEE Pervasive Computing. 2004;**3**(4):11-15

[43] Mohammadzadeh N, Safdari R. Patient monitoring in mobile health: Opportunities and challenges. Medieval Archaeology. 2014;**68**(1):57

[44] Alnosayan N, Lee E, Alluhaidan A, Chatterjee S, Houston-Feenstra L, Kagoda M, Dysinger W. MyHeart: An intelligent mhealth home monitoring system supporting heart failure self-care. In: 2014 IEEE 16th International Conference on e-Health Networking, Application & Services (HealthCom); 2015. pp. 311-316

[45] Minutolo A, Sannino G, Esposito M, De Pietro G. A rule-based mHealth system for cardiac monitoring. In: Proceedings of 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES); 2010. pp. 144-149

[46] Boulos MNK, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: An overview, with example from eCAALYX. Biomedical Engineering Online. 2011;**10**:1-14

[47] Islam SMR, Kwak D, Kabir H, Hossain M, Kwak K-S. The Internet of things for health care: A comprehensive survey. IEEE Access. 2015;**3**:678-708

[48] Da Xu L, He W, Li S. Internet of things in industries: A survey. IEEE Transactions on Industrial Informatics. 2014;**10**(4):2233-2243

[49] Costa R, Carneiro D, Novais P, Lima L, Machado J, Marques A, Neves J. Ambient assisted living. Applied Soft Computing. 2008;**51**:86-94

[50] Amiribesheli M, Benmansour A, Bouchachia A. A review of smart homes in healthcare. Journal of Ambient Intelligence and Humanized Computing. 2015;**6**(4):495-517

[51] Pollack ME, Engberg S, Matthews JT, Dunbar-jacob J, Mccarthy CE, Thrun S. Pearl: A mobile robotic assistant for the elderly. Architecture. 2002;**2002**:85-91

[52] Chen Y, Argentinis E, Weber G. IBM Watson: How cognitive computing can be applied to big data challenges in life sciences research. Clinical Therapeutics. 2016;**38**(4):688-701

[53] Devarakonda M, Tsou C-H. Automated problem list generation from electronic medical records in IBM Watson. In: 29th AAAI Conference on Artificial Intelligence (AAAI) 2015 and the 27th Innovative Applications of Artificial Intelligence Conference, IAAI 2015; vol. 5; 2015. pp. 3942-3947

Chapter 7

Phoebe Framework and Experimental Results for Estimating Fetal Age and Weight

Loc Nguyen, Truong-Duyet Phan and
Thu-Hang T. Ho

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.74883

Abstract

Fetal age and weight estimation plays an important role in pregnant treatments. There are many estimation formulas created by the combination of statistics and obstetrics. However, such formulas give optimal estimation if and only if they are applied into specified community. This research proposes a so-called Phoebe framework that supports physicians and scientists to find out most accurate formulas with regard to the community where scientists do their research. The built-in algorithm of Phoebe framework uses statistical regression technique for fetal age and weight estimation based on fetal ultrasound measures such as bi-parietal diameter, head circumference, abdominal circumference, fetal length, arm volume, and thigh volume. This algorithm is based on heuristic assumptions, which aim to produce good estimation formulas as fast as possible. From experimental results, the framework produces optimal formulas with high adequacy and accuracy. Moreover, the framework gives facilities to physicians and scientists for exploiting useful statistical information under pregnant data. Phoebe framework is a computer software available at http://phoebe.locnguyen.net.

Keywords: fetal age estimation, fetal weight estimation, ultrasound measures, regression model, estimation formula

1. Introduction

Fetal age and weight estimation is to predict the birth weight or birth age before delivery. It is very important for doctors to diagnose abnormal or diseased cases so that she/he can decide treatments on such cases. Because this research mentions both age estimation and weight

estimation, for convenience, the term "birth estimation" implicates both of them. There are two methods for birth estimation:

- Determining volume of fetal inside mother womb and then calculating fetal weight based on such volume and mass density of flesh and bone. By the other way, fetal age and weight can be estimated according to size of mother womb.
- Applying statistical regression model: Fetal ultrasound measures such as bi-parietal diameter (*bpd*), head circumference (*hc*), abdominal circumference (*ac*), fetal length (*fl*), arm volume (*arm_vol*), and thigh volume (*thigh_vol*) are recorded and considered as input sample for regression analysis which results in a *regression function*. This function is formula for estimating fetal age and weight according to ultrasound measures such as *bpd*, *hc*, *ac*, *fl*, *arm_vol*, and *thigh_vol*. Data that are composed of these ultrasound measures are called gestational sample or pregnant sample. Terms: "*sample*" and "*data*" have the same meaning in this research. Sample is representation of population where research takes place.

Because the second method reflects features of population from statistical data, the regression model is chosen for birth estimation in this research. Note, some terminologies such as *function*, *regression function*, *estimation function*, *regression model*, *estimation model*, *formula*, *regression formula*, and *estimation formula* have the same meaning.

There are many estimation formulas resulted from gestational researches such as [1–9]. Some of them gain high accuracy, but they are only appropriate to population, community or ethnic group, where such researches are done. If we apply these formulas into other community such as Vietnam, they are no longer accurate. Moreover, it is difficult to find out a new and effective estimation formula or the cost of time and (computer) resources of formula discovery is expensive. Therefore, the first goal of this research is to propose an effective built-in algorithm, which produces highly accurate formulas that are easy to tune with specified population. The process of producing formulas by such algorithm is as fast as possible. In addition, physicians and researchers always want to discover useful statistical information from measure sample and regression model. Thus, the second goal of this research is to give facilities to physicians and researchers by introducing them a framework that is called *Phoebe framework* or *Phoebe system*. Phoebe framework implements such built-in algorithm in the first goal and provides a tool allowing physicians and researchers to exploit and take advantage of useful information under gestational sample. This tool is programmed as computer software. Moreover, Phoebe framework allows software developers to modify its modules. For example, developers can improve the built-in algorithm by adding heuristic constraints.

This chapter is the improved collection of our two articles "A framework of fetal age and weight estimation" [10] and "Experimental Results of Phoebe Framework: Optimal Formulas for Estimating Fetus Weight and Age" [11]. Section 2 gives an overview of the architecture of Phoebe framework. Section 3 is a description of the built-in algorithm to produce optimal formulas which are appropriated to a concrete population like Vietnam. Such algorithm is the core of Phoebe framework. Section 4 discusses main use cases of the framework with respect to gestational sample. As experimental results, some interesting estimation formulas produced by the framework are described in Section 5. A proposal of early weight estimation is proposed in Section 6. Conclusion is given in Section 7. Note that Phoebe framework used statistic software

package "Java Scientific Library" of Michael Thomas Flanagan [12] and parsing package "A Java expression parser" of Jos de Jong [13]. The package "Java Scientific Library" is the most important one in the framework. The framework is implemented by Java language [14].

2. General architecture of Phoebe framework

Based on clinical data input which includes fetal ultrasound measures such as *bpd*, *hc*, *ac*, and *fl*, the framework produces optimal formulas for estimating fetal weight and fetal age with the highest precision. Moreover, statistical information about fetus and gestation is also described in detail with two forms: numerical format and graph format. Therefore, the framework consists of four components as follows:

- *Dataset* component is responsible for managing information about fetal ultrasound measures such as *bpd*, *hc*, *ac*, *fl* and extra gestational information in reasonable and intelligent manner. This component allows other components to retrieve such information. Gestational information is organized into some abstract structure, for example, a matrix, where each row represents a sample of *bpd*, *hc*, *ac*, *fl* measures. **Table 1** is an example of this abstract structure.
- *Regression* component represents estimation formula or regression function. This component reads ultrasound information from *Dataset* component and builds up optimal estimation formula from such information. The built-in algorithm, which is used to discover and construct estimation formula, is discussed in Section 3. This component is the most important one because it implements such discovery algorithm.
- *Statistical Manifest* component describes statistical information of both ultrasound measures and regression function, for example, mean and standard deviation of *bpd* samples, sum of residuals, correlation coefficient of regression function, and percentile graph of

bpd	hc	fl	ac	Fetal age (week)	Fetal weight (gram)
74	262	51	255	28	900
72	260	51	232	28	900
68	260	50	229	28	900
72	275	52	240	28	900
72	274	52	240	28	950
74	253	50	235	28	950
71	257	52	239	28	950
71	255	53	236	28	950
70	264	52	246	28	950

Table 1. An example of gestational sample matrix.

fetal weight. Statistical manifest is organized into two forms such as numerical format and graph format.

- *User Interface* (*UI*) component is responsible for providing interaction between system and users such as physicians and researchers. A popular use case is that users enter ultrasound measures and require system to print out both optimal estimation formula and statistical information about such ultrasound measures; moreover, users can retrieve other information in *Dataset* component. *UI* component links to all of other components so as to give users as many facilities as possible.

Three components: *Dataset*, *Regression* and *Statistical Manifest* are basic components. The fourth component *User Interface* is the bridge among them. **Figure 1** shows a general architecture of Phoebe framework.

3. Built-in algorithm of Phoebe framework

Phoebe framework uses a regression model for estimating fetal weight and age. Suppose a linear regression function $Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_n X_n$ where Y is fetal weight or age, whereas X_i (s) are gestational ultrasound measures such as *bpd*, *hc*, *ac*, and *fl*. Variable Y is called response variable or dependent variable. Each X_i is called *regression variable*, *regressor*, *regression variable*, or *independent variable*. Each α_i is called *regression coefficient*. Given a set of measure values of X_i (s), the value of Y called *Y-estimated* calculated from this regression function is estimated fetal weight (or age) which is compared with real value of Y measured from ultrasonic machine. The real value of Y called *Y-real* is fetal weight (or age) available in sample. In this research, the notation Y refers implicitly to *Y-estimated* if there is no explanation. The deviation between *Y-estimated* and *Y-real* is a criterion used to assess the quality or the precision of regression function. This deviation is also called *estimation error*. The less the deviation is, the better the regression function is. The goal of this research is to find out the optimal regression function or estimation formula whose precision is highest.

A regression function will be good if it meets two conditions as follows:

- The correlation between *Y-estimated* and *Y-real* is large.

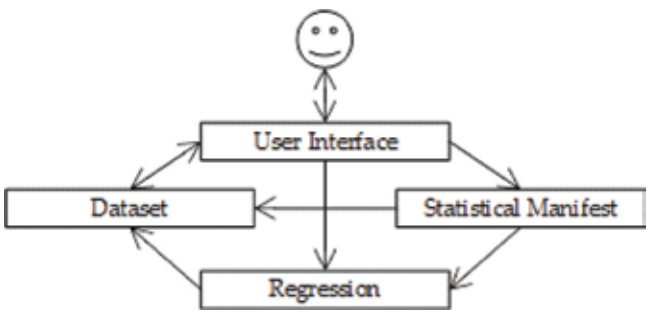

Figure 1. General architecture of Phoebe framework.

- The sum of residuals is small. Note that residual is defined as the square of deviation between *Y-estimated* and *Y-real*. We have:

$$residual = (Y_{estimated} - Y_{real})^{2.}$$

These two conditions are called the *pair of optimal conditions*. A regression function is optimal or best if it satisfies the pair of optimal conditions at most, where correlation between *Y-estimated* and *Y-real* is largest, and the sum of residuals is smallest. Given a set of regression variables X_i (where $i = 1, 2,\ldots, n$), we recognize that a regression function is a combination of k variables X_i (s) where $k \leq n$ so that such combination achieves the pair of optimal conditions. Given a set of possible regression variables $VAR = \{X_1, X_2,\ldots, X_n\}$ being ultrasound measures, brute-force algorithm can be used to find out optimal function, which includes three following steps:

1. Let indicator number k be initialized 1, which responds to k-combination having k regression variables.
2. All combinations of n variables taken k are created. For each k-combination, the function built up by k variables in this k-combination is evaluated on the pair of optimal conditions; if such function satisfies these conditions at most then, it is optimal function.
3. Indicator k is increased by 1. If $k = n$ then algorithm stops, otherwise go back step 2.

The number of combinations which brute-force algorithm browses is:

$$\sum_{k=1}^{n} \frac{n!}{k!(n-k)!}$$

where n is the number of regression variables and notation, and "$k!$" denotes factorial of k. If n is large enough, there are a huge number of combinations, which causes that the brute-force algorithm never terminates and it is impossible to find out the best function. Moreover, there are many kinds of regression function such as linear, quadric, cube, logarithm, exponent, and product. Therefore, we propose an algorithm which overcomes this drawback and always finds out the optimal function. In other words, the termination of the proposed algorithm is determined, and the time cost is decreased significantly because the searching space is reduced as small as possible. The proposed algorithm is called *seed germination* (***SG***) algorithm. SG is built-in algorithm of Phoebe framework, which is the core of Phoebe framework. It is heuristic algorithm, which is based on the *pair of heuristic assumptions* as follows:

- First assumption: regression variables X_i (s) trends to be mutually independent. It means that any pair of X_i and X_j with $i \neq j$ in an optimal function are mutually independent. The independence is reduced into the looser condition "*the correlation coefficient of any pair of X_i and X_j is less than a threshold δ.*" This is *minimum* assumption.
- Second assumption: each variable X_i contributes to quality of optimal function. The contribution rate of a variable X_i is defined as the correlation coefficient between such variable and *Y-real*. The higher the contribution rate is, the more important the respective variable is. Variables with high contribution rate are called *contributive* variables. Therefore, optimal

function includes only contributive regression variables. The second assumption is stated that "*the correlation coefficient of any regression variable X_i and real response value Y-real is greater than a threshold ε*." This is the *maximum* assumption.

SG algorithm tries to find out a combination of regression variables X_i (s) so that such combination satisfies such pair of heuristic assumptions. In other words, it is expected that this combination can constitute an optimal regression function that satisfies the *pair of heuristic conditions*, as follows ([10] p. 22):

- The correlation coefficient of any pair of X_i and X_j is less than the minimum threshold $\delta > 0$. This condition is corresponding to the minimum assumption, which is called *minimum condition* or *independence condition*.
- The correlation coefficient of any X_i and *Y-real* is greater than the maximum threshold $\varepsilon > 0$. This condition is corresponding to the maximum assumption, which is called *maximum condition* or *contribution condition*.

Given a set of possible regression variables $VAR = \{X_1, X_2, \ldots, X_n\}$ being ultrasound measures, let $f = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_k X_k$ $(k \leq n)$ be the estimation function and let $Re(f) = \{X_1, X_2, \ldots, X_k\}$ be its regression variables. Note that the value of f is fetal age or fetal weight. $Re(f)$ is considered as the representation of f. Let *OPTIMAL* be the output of SG algorithm, which is a set of optimal functions returned. *OPTIMAL* is initialized as empty set. Let $Re(OPTIMAL)$ be a set of regression variables contained in all optimal functions $f \in OPTIMAL$. SG algorithm has four following steps ([10] p. 22):

1. Let C be the complement set of *VAR* with regard to *OPTIMAL*, we have $C = \text{VAR} \backslash Re(OPTIMAL)$ where the backslash "\" denotes complement operator in set theory. It means that C is in *VAR* but not in $Re(OPTIMAL)$.

2. Let $G \subset C$ be a list of regression variables satisfying the pair of heuristic conditions. Note, G is subset of C. If G is empty, the algorithm terminates; otherwise going to step 3.

3. We iterate over G in order to find out the candidate list of good functions. For each regression variable $X \in G$, let L be the union set of optimal regression variables and X. We have $L = Re(f) \cup \{X\}$ where $f \in OPTIMAL$. Suppose *CANDIDATE* is a candidate list of good functions, which is initialized as empty set. Let g be the new function created from L; in other words, regression variables of g belong to L, $Re(g) = L$. If function g meets the pair of heuristic conditions, it is added into *CANDIDATE*, $CANDIDATE = CANDIDATE \cup \{g\}$.

4. Let *BEST* be a set of best functions taken from *CANDIDATE*. In other words, these functions belong to *CANDIDATE* and satisfy the pair of heuristic conditions at most, where correlation is the largest and the sum of residuals is the smallest. If *BEST* equals *OPTIMAL*, then the algorithm stops; otherwise assigning *BEST* to *OPTIMAL* and going back step 1. Note that two sets are equal if their elements are the same.

Figure 2 shows the flow chart of SG algorithm.

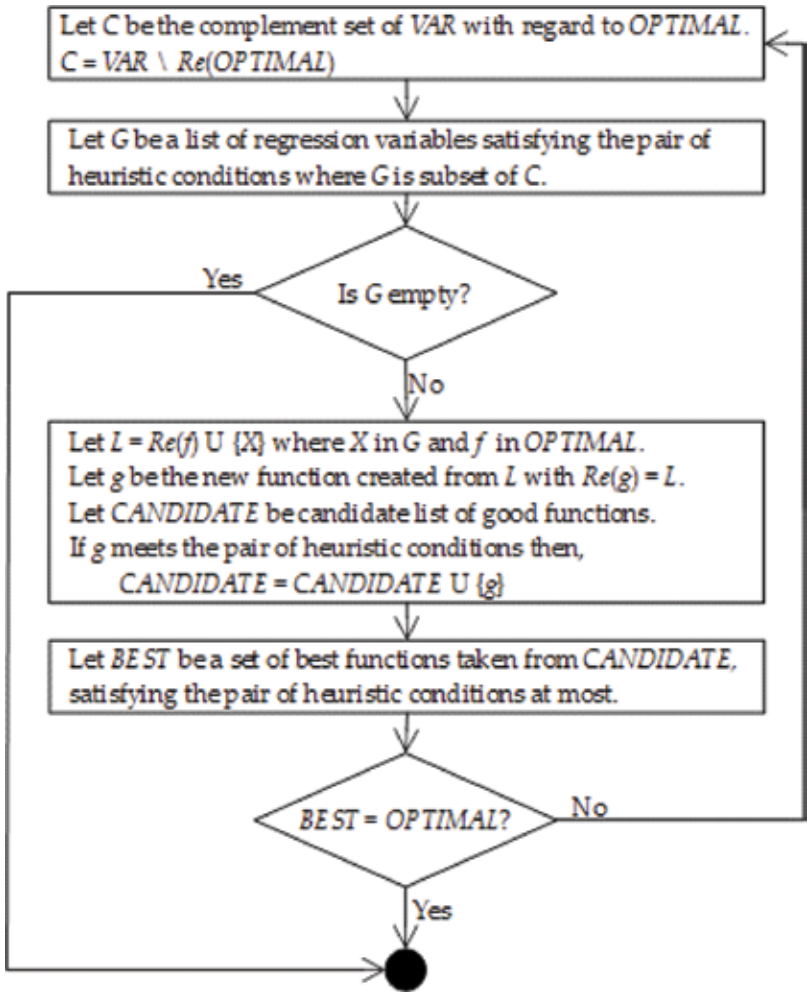

Figure 2. Flow chart of SG algorithm.

SG algorithm was described in article "A framework of fetal age and weight estimation" ([10] pp. 21–23). It is easy to recognize that the essence of SG algorithm is to reduce search space by choosing regression variables satisfying heuristic assumption as "seeds." Optimal functions are composed of these seeds. The algorithm always delivers best functions but can lose other good functions. The length of function is defined as the number of its regression variables. Terminated condition is that no more optimal functions can be found out or possible variables are browsed exhaustedly. Therefore, the result function is the longest and best one, but some other shorter functions may be significantly good.

The current implementation of SG algorithm establishes that the minimum threshold δ is arbitrary. It also supports nonlinear regression models shown in **Table 2** as follows:

Polynomial	$Y = \alpha_0 + \alpha_1(X_1 + X_2 + \ldots + X_n)^k$
Logarithm	$Y = \alpha_0 + \alpha_1\log(X_1) + \alpha_2\log(X_2) + \ldots + \alpha_n\log(X_n)$
	$Y = \alpha_0 + \alpha_1\log(X_1 + X_2 + \ldots + X_n)$
Exponent	$Y = \exp(\alpha_0 + \alpha_1X_1 + \alpha_2X_2 + \ldots + \alpha_nX_n)$
	$Y = \exp(\alpha_0 + \alpha_1(X_1 + X_2 + \ldots + X_n))$
Product	$Y = \alpha_0{X_1}^{\alpha_1}{X_2}^{\alpha_2}\ldots{X_n}^{\alpha_n}$

Table 2. Nonlinear regression models.

The notations "exp" and "log" denote exponent function and natural logarithm function, respectively. Most of nonlinear regression models can be transformed into linear regression models. For example, given the product model, the following is an example of linear transformation.

$$\log(Y) = \log(\alpha_0) + \alpha_1\log(X_1) + \alpha_2\log(X_2) + \ldots + \alpha_n\log(X_n)$$

Let,

$$U = \log(Y), Z_i = \log(X_i), \beta_0 = \log(\alpha_0), \beta_{i\geq 1} = \alpha_i$$

The product model becomes the linear model with regard to variables U, Z_i and coefficients β_i as follows:

Polynomial transformation	$Y = \alpha_0 + \alpha_1(X_1 + X_2 + \ldots + X_n)^k$
	$Y = \alpha_0 + \alpha_1Z_1$ where $Z_1 = (X_1 + X_2 + \ldots + X_n)^k$
Logarithm transformation	$Y = \alpha_0 + \alpha_1\log(X_1) + \alpha_2\log(X_2) + \ldots + \alpha_n\log(X_n)$
	$Y = \alpha_0 + \alpha_1Z_1 + \alpha_2Z_2 + \ldots + \alpha_nZ_n$ where $Z_i = \log(X_i)$
Logarithm transformation	$Y = \alpha_0 + \alpha_1\log(X_1 + X_2 + \ldots + X_n)$
	$Y = \alpha_0 + \alpha_1Z_1$ where $Z_1 = \log(X_1 + X_2 + \ldots + X_n)$
Exponent transformation	$Y = \exp(\alpha_0 + \alpha_1X_1 + \alpha_2X_2 + \ldots + \alpha_nX_n)$
	$U = \alpha_0 + \alpha_1X_1 + \alpha_2X_2 + \ldots + \alpha_nX_n$ where $U = \log(Y)$
Exponent transformation	$Y = \exp(\alpha_0 + \alpha_1(X_1 + X_2 + \ldots + X_n))$
	$U = \alpha_0 + \alpha_1Z_1$ where $U = \log(Y)$ and $Z_1 = X_1 + X_2 + \ldots + X_n$
Product transformation	$Y = \alpha_0{X_1}^{\alpha_1}{X_2}^{\alpha_2}\ldots{X_n}^{\alpha_n}$
	$U = \beta_0 + \beta_1Z_1 + \beta_2Z_2 + \ldots + \beta_nZ_n$ where $U = \log(Y), Z_i = \log(X_i), \beta_0 = \log(\alpha_0), \beta_{i\geq 1} = \alpha_i$

Table 3. Transformation of nonlinear models into linear models.

$$U = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \ldots + \beta_n Z_n$$

Table 3 shows how to transform nonlinear models into linear models.

With the built-in SG algorithm, Phoebe framework can be totally used for any regression application beyond birth estimation.

4. Use cases of Phoebe framework

Phoebe framework has three basic use cases realized by three components: dataset, regression model and statistical manifest as discussed in Section 2. Three basic use cases include:

1. Discovering optimal formulas with high accuracy. Optimal formulas are results of SG algorithm described in Section 3.
2. Providing statistical information under gestational sample. Statistical information is in numeric format and graph format.
3. Comparison among different formulas.

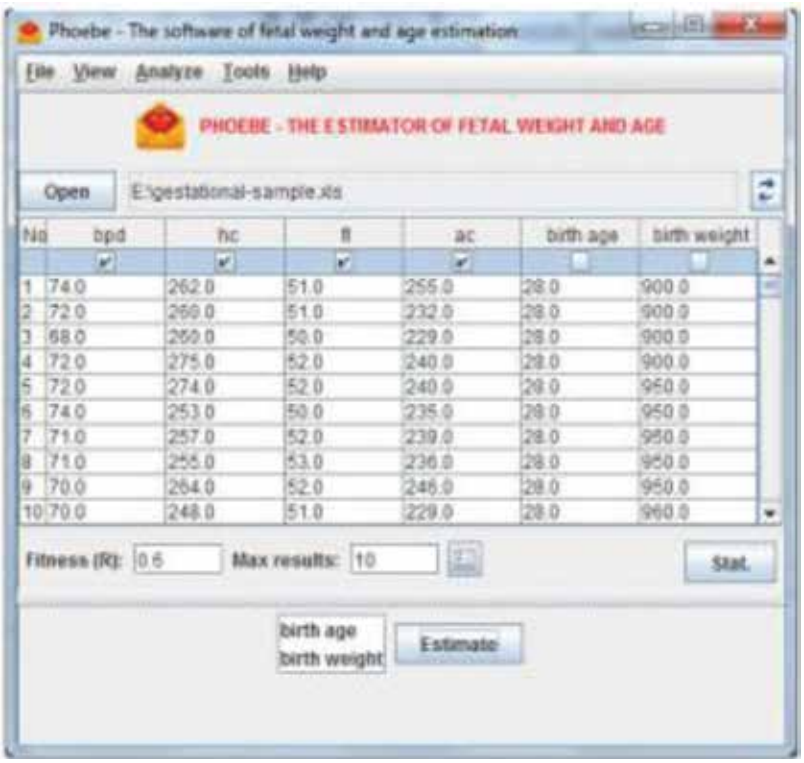

Figure 3. Gestational sample.

Use case 1: Discovering optimal formulas

Given gestational data [15] are composed of two-dimensional ultrasound measures of pregnant women. These measures are taken at Vinh Long General Hospital – Vietnam, which include bi-parietal diameter (*bpd*), head circumference (*hc*), abdominal circumference (*ac*) and fetal length (*fl*). Fetal age is from 28 to 42 weeks. Fetal weight is measured by gram. Gestational sample is shown in **Figure 3**.

After specifying the maximum threshold ε (fitness value) and which measures are regression variables and response variable, user presses button "Estimate" to retrieve optimal formulas as results of SG algorithm. Such optimal formulas are shown in **Figure 4**. Note, in **Figure 4**, regression variables are *bpd*, *hc*, *ac*, and *fl*, whereas response variable is fetal weight. The threshold ε is 0.6.

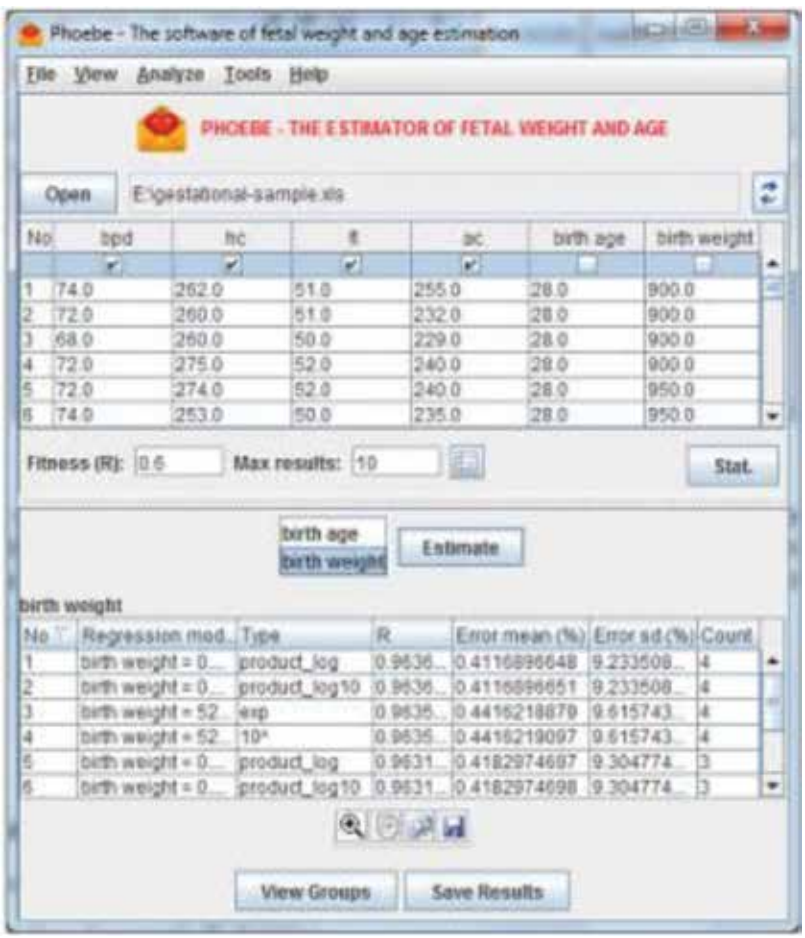

Figure 4. Optimal weight estimation formulas.

http://dx.doi.org/10.5772/intechopen.74883

An estimation formula with one or two regressors (ultrasound measures) can be represented as a graph. In the illustrative **Figure 5**, the horizontal axis indicates the measure *bpd* in millimeter, and the right vertical axis indicates the measure *ac* in millimeter. The left vertical axis shows the estimated weight.

The graph in **Figure 5** has 11 estimation lines represented as internal (red) lines. Each estimation line corresponds to a small interval of *ac*. Fetal weight on each estimation line ranges from 900 to 4800 g. This is a way to show a three-dimensional function as a two-dimensional graph. For example, given *bpd* = 90 and *ac* = 300, we need to estimate fetal weight. Because *ac* is 300 mm, we look at the sixth estimation line from bottom to up. The intersection point between *bpd* = 90 and the sixth estimation line is projected on the left vertical axis, which results out a fetal weight that approximates to (4800–900)/2 + 900 ≈ 2850 g because such intersection point is near to midpoint of the weight range on the sixth estimation line.

Use case 2: Providing statistical information

Statistical information is classified into two groups: gestational information and estimation information.

- Gestational information contains statistical attributes about fetal ultrasound measures, for example, mean, median and standard deviation of *bpd*.
- Estimation information contains attributes about estimation model, for example, correlation coefficient, sum of residuals and estimation error of estimation formula.

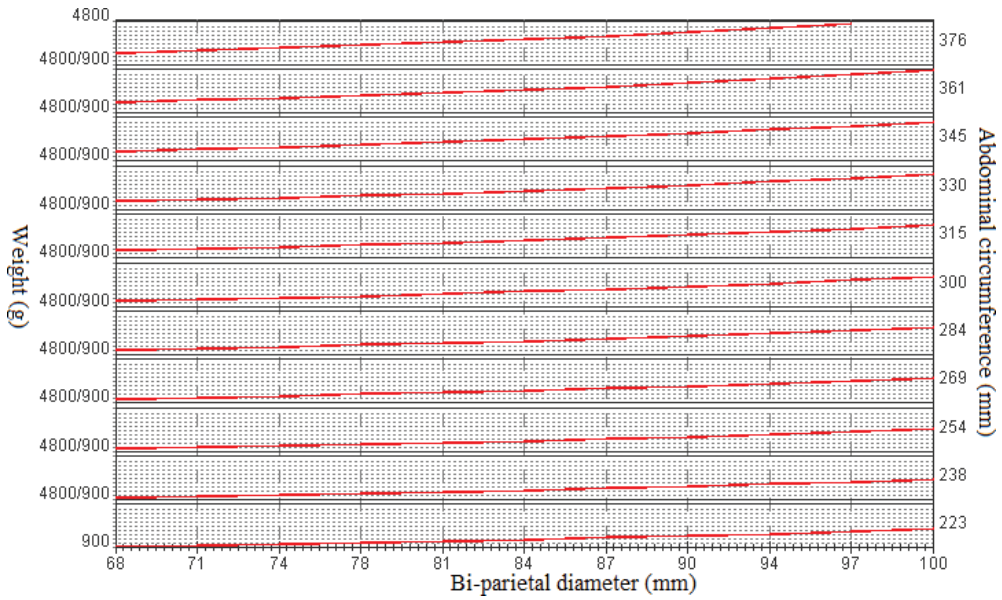

Figure 5. Estimation graph for estimating fetal weight.

In representation, statistical information is described in two forms: numeric format and graph format. **Figure 6** shows statistical attributes (mean, median, standard deviation, histogram, etc.) of fetal age and ultrasound measures *bpd*, *hc*, *ac*, *fl*.

Figure 7 shows a full description of a weight estimation formula: *weight* = 0.000043 * (*bpd*^1.948640) * (*hc*^0.263745) * (*fl*^0.601972) * (*ac*^0.905524). For instance, sum of residuals (SS) is 46412446.0047 and estimation error is -7.4655 ± 212.5571. Note, the sign "^" denotes exponent function, for example, 2^3 = 8.

Use case 3: Comparison among different formulas

There are many criteria to evaluate efficiency and accuracy of estimation formulas. These criteria are called evaluation criteria, for example, correlation coefficient, sum of residuals, estimation error. Each formula has individual strong points and drawbacks. A formula is better than another one in terms of some criteria but may be worse than this other one in terms of different criteria. An optimal formula is the one that has more strong points than drawbacks

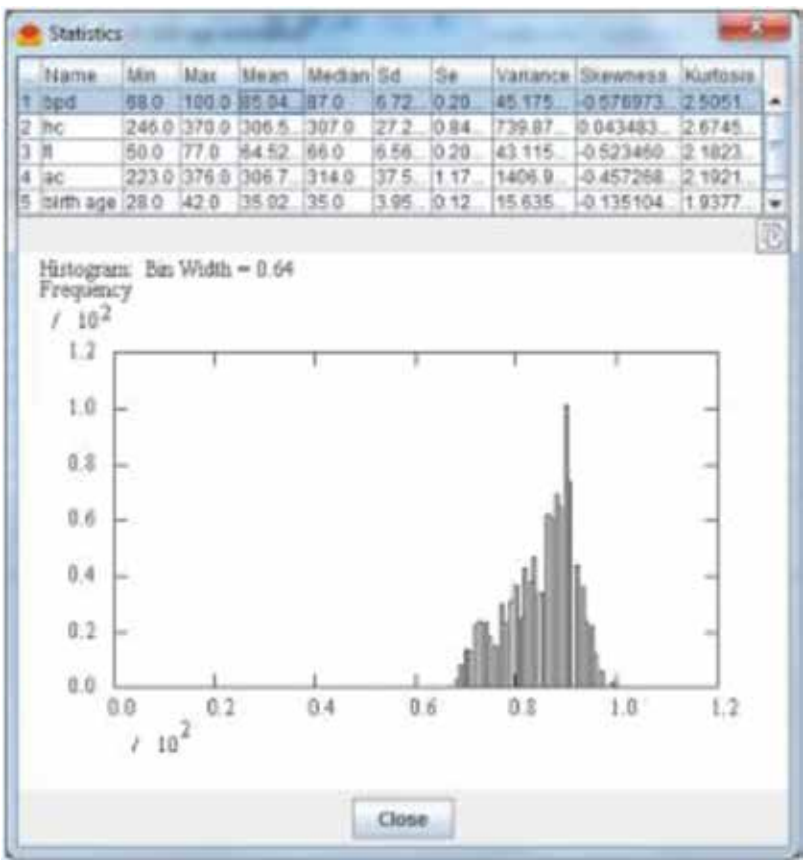

Figure 6. Gestational statistical information.

http://dx.doi.org/10.5772/intechopen.74883

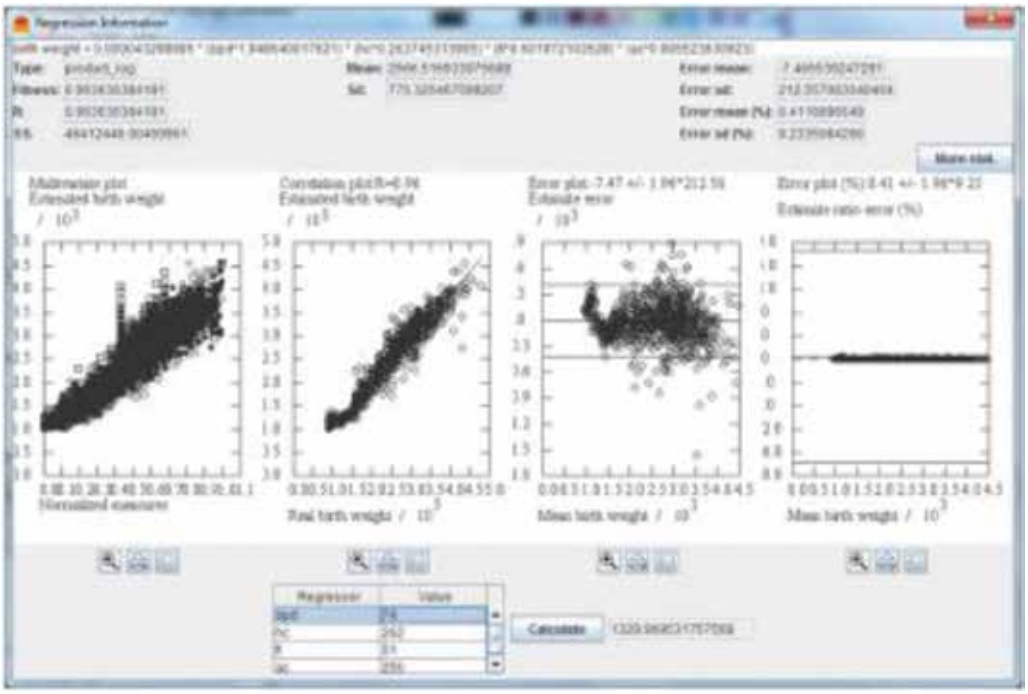

Figure 7. Statistical estimation information.

in most criteria. Hence, Phoebe framework supports the comparison among different formulas via *evaluation matrix* represented in **Figure 8**. Each row in evaluation matrix represents a formula, whereas each column indicates a criterion. For example, first row, second row and third row represent three formulas in form of logarithm function, exponent function and linear function, respectively. Four criteria such as multivariate correlation, estimation correlation, error range and ratio error range are arranged in three respective columns.

Tables 4–8 in the section "experimental results" are numeric interpretations of evaluation matrix in **Figure 8**.

5. Experimental results

We make experiments based on Phoebe framework in order to find out optimal formulas for estimating fetus weight and age with note that such formulas are most appropriate to our gestational samples. We use two samples in which the first sample includes two-dimensional (2D) ultrasound measures of 1027 cases and the second sample includes three-dimensional (3D) ultrasound measures of 506 cases. Ho and Phan [15, 16] collected these samples of pregnant women at Vinh Long General Hospital, Vietnam, with obeying strictly all medical ethical criteria. These women and their husbands are Vietnamese. Their periods are regular,

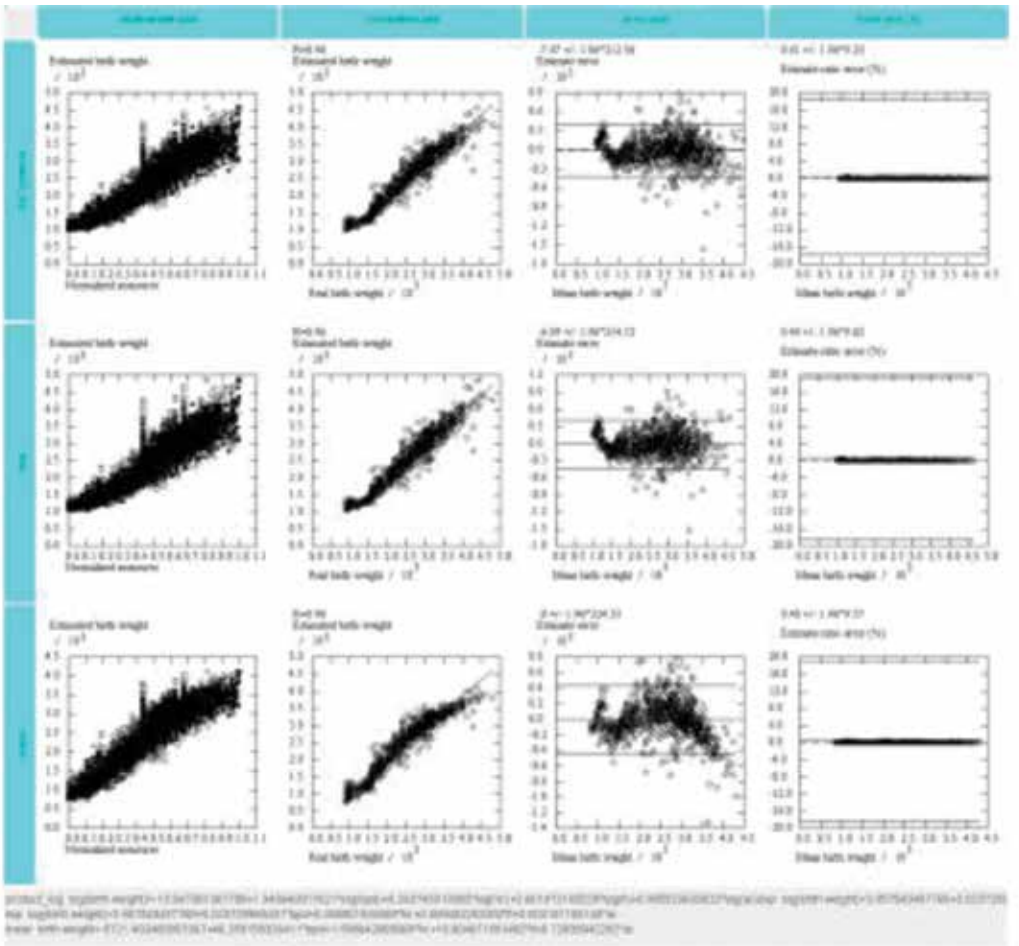

Figure 8. Comparison among different formulas.

and their last periods are determined. Each of them has only one alive fetus. Fetal age is from 28 to 42 weeks. Delivery time is not over 48 h since ultrasound scan. Measures in 2D sample are *bpd*, *hc*, *ac*, and *fl*. Measures in 3D sample are *bpd*, *hc*, *ac*, *fl*, *thigh_vol*, *arm_vol*. The unit of *bpd*, *hc*, *ac*, *fl* is millimeter. The unit of *thigh_vol* and *arm_vol* is cm^3. The units of fetal age and fetal weight are week and gram, respectively. Experimental results mentioned in this section were also published in our article "Experimental Results of Phoebe Framework: Optimal Formulas for Estimating Fetus Weight and Age" [11].

The proposed framework can produce amazing formulas. We compare our optimal formulas with the others according to metrics such as estimation correlation and estimation error range, given such two gestational samples. Let $Y = \{y_1, y_2, \ldots, y_n\}$ and $Z = \{z_1, z_2, \ldots, z_n\}$ be fetal sample age/weight and fetal estimated age/weight, respectively. The estimation correlation denoted R is correlation coefficient of sample response value and estimated response value, according to Eq. (1). The correlation R reflects adequacy of a given formula. The larger the R is, the better the formula is:

$$R = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(z_i - \bar{z})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(z_i - \bar{z})^2}} \quad (1)$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$$

An estimation error denoted d_i is deviation between z_i and y_i. The estimation error mean denoted μ is mean of errors. The error mean μ reflects accuracy of a given formula. The smaller the absolute value of μ is, the more accurate the formula is. If μ is positive, the respective formula leans to overestimation. If μ is negative, the respective formula leans to low estimation. The standard deviation σ of estimation errors reflects the stability of a given formula. The smaller the standard deviation σ is, the more stable the formula is. The combination of error mean μ and standard deviation σ results out a so-called *error range*. Eq. (2) explains how to calculate μ, σ, and error range.

$$\begin{aligned} d_i &= z_i - y_i \\ \mu &= \frac{1}{n}\sum_{i=1}^{n} d_i \\ \sigma &= \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(d_i - \mu)^2} \\ error\_range &= [\mu - \sigma, \mu + \sigma] = \mu \pm \sigma \end{aligned} \quad (2)$$

For example, if $\mu = -0.0292$ and $\sigma = 1.45$ then, the error range is -0.0292 ± 1.45, which means that the total average error ranges from $-1.4792 = -0.0292\text{-}1.45$ to $1.4208 = -0.0292 + 1.45$. The error range reflects both adequacy and accuracy of a given formula.

Formula	Expression	R	Error range
NH 1	log(*age*) = 2.419638 + 0.002012 * *bpd* + 0.000934 * *hc* + 0.00547 * *fl* + 0.001042 * *ac*	0.9303	−0.0292 ± 1.4500
NH 2	*age* = −3.364759 + 0.056285 * *bpd* + 0.034697 * *hc* + 0.188156 * *fl* + 0.035304 * *ac*	0.9285	0 ± 1.4682
Ho 1	*age* = 331.022308–1.611774 * (*hc* + *ac*) + 0.00278 * ((*hc* + *ac*)^2) - 0.000002 * ((*hc* + *ac*)^3)	0.9212	0 ± 1.5384
Varol 6	*age* = 11.769 + 1.275 * *fl*/10 + 0.449 * ((*fl*/10)^2) - 0.02 * ((*fl*/10)^3)	0.8949	−1.6807 ± 1.8525
Varol 1	*age* = 5.596 + 0.941 * *ac*/10	0.8941	−0.5683 ± 1.7711
Varol 5	*age* = 1.863 + 6.280 * *fl*/10–0.211 * ((*fl*/10)^2)	0.8934	−1.5182 ± 2.1150

The sign "^" denotes exponent operator. The template of formulas aims to flexibility, which can be input of any computational tool. **Table 5** shows a comparison between our best weight formula and the others with 2D sample. As seen in **Table 5**, our formula is the best with $R = 0.9636$ and error range − 7.4656 ± 212.5573 g.

Table 4. Comparison of age estimation with 2D sample.

Table 4 shows a comparison between our best age formula and the others with 2D sample. As a convention, the name of each formula is the name of respective author listed in references section. For example, formula "Ho 1" is the first formula of the author Ho [4]. As seen in **Table 4**, our formula is the best with $R = 0.9303$ and error range -0.0292 ± 1.4500 week (s). As a convention, our formulas have names with prefix "NH"

Table 6 shows comparison between our best age formula and the others with 3D sample. As seen in **Table 6**, our formula is the best with $R = 0.9970$ and error range ± 0.2696 week

Table 7 shows a comparison between our best weight formula and the others with 3D sample. As seen in **Table 7**, our formula is the best with $R = 0.9708$ and error range -0.0001 ± 180.9803 g

Within the context of this research, from section of 3D ultrasound in PhD dissertation of Ho [4], I recognize that fetus weight and fetus age are mutually dependent. For instance, when fetus age increases, fetus weight increases too. As a result, weight estimation is improved significantly if fetus age was known before. If fetus age is added into the regression model of fetus weight as a regression variable (regressor), the resulted weight estimation formula, called *dual formula*, is even better than the most optimal ones shown in **Tables 5** and **6**. Such dual formula is not only precise but also practical because many pregnant women knew their gestational age before taking an ultrasound examination. Given 2D sample and 3D sample, **Table 8** shows dual formulas in comparison with the most optimal ones shown in **Tables 5** and **7** with regard to *R* and error range. As a convention, our dual formulas have names with prefix "NHD". Notation "log10" denotes logarithm function with base 10.

Formula	Expression	R	Error range
NH 3	log(*weight*) = −10.047381 + 1.94864 * log(*bpd*) + 0.263745 * log (*hc*) + 0.601972 * log(*fl*) + 0.905524 * log(*ac*)	0.9636	−7.4656 ± 212.5573
NH 4	log(*weight*) = 3.957543 + 0.02373 * *bpd* + 0.000802 * *hc* + 0.009403 * *fl* + 0.003157 * *ac*	0.9635	−6.0901 ± 214.1153
Sherpard	*weight* = 10^(1.2508 + 0.166 * *bpd*/10 + 0.046 * *ac*/10–0.002646 * *ac* * *bpd*/100)	0.9619	−65.8121 ± 219.0392
Ho 2	*weight* = 10^(1.746 + 0.0124 * *bpd* + 0.001906 * *ac*)	0.9602	−11.5576 ± 223.5124
Hadlock	*weight* = 10^(1.304 + 0.05281 * *ac*/10 + 0.1938 * *fl*/10–0.004 * *ac* * *fl*/100)	0.9395	−76.4960 ± 272.9474
Campbell and Wilkin	*weight* = 1000 * exp.(−4.564 + 0.282 * *ac*/10–0.00331 * *ac* * *ac*/100)	0.9215	68.1261 ± 308.5728

Table 5. Comparison of weight estimation with 2D sample.

Formula	Expression	R	Error range
NH 5	*age* = 20.759763 + 0.170859 * (*thigh_vol* + *arm_vol*) - 0.000545 * ((*thigh_vol* + *arm_vol*) ^2) + 0.000001 * ((*thigh_vol* + *arm_vol*)^3)	0.9970	0 ± 0.2696
NH 6	*age* = 21.816252 + 0.137531 * (*thigh_vol* + *arm_vol*) - 0.000228 * ((*thigh_vol* + *arm_vol*) ^2)	0.9969	0 ± 0.2752
Ho 3	*age* = 21.1148 + 0.2381 * *thigh_vol* - 0.001 * (*thigh_vol*^2) + 0.000002 * (*thigh_vol*^3)	0.9960	−0.0150 ± 0.3173
Ho 4	*age* = 167.079079–1.553705 * *ac* + 0.005559 * (*ac*^2) - 0.000006 * (*ac*^3)	0.8482	0.3723 ± 1.8985

Table 6. Comparison of age estimation with 3D sample.

In **Table 8**, all dual formulas NHD * are better than normal formulas NH * with regard to *R* and error range. Moreover, NHD * do not need too much regressors. Given 2D sample, NHD 1 and NHD 2 use 4 and 3 regressors including age regressor, respectively whereas both NH 3 and NH 4 uses 4 regressors. Given 3D sample, NHD 3 and NHD 4 use 6 and 5 regressors including age regressor, respectively, whereas NH 7 and NH 8 use 5 and 3 regressors, respectively.

Formula	Expression	R	Error range
NH 7	*weight* = −3617.936175 + 0.513171 * *hc* + 1.960176 * *ac* + 39.804645 * *bpd* + 17.016936 * *fl* + 8.366404 * *thigh_vol* + 5.828808 * *arm_vol*	0.9708	−0.0001 ± 180.9803
NH 8	*weight* = −3626.314419 + 43.426744 * *bpd* + 23.645338 * *fl* + 11.414273 * *thigh_vol*	0.9698	0 ± 184.0439
Ho 5	*weight* = −3306 + 55.477 * *bpd* + 13.483 * *thigh_vol*	0.9663	−0.0072 ± 194.0956
Lee 3	*weight* = exp.(0.5046 + 1.9665 * log(*bpd*/10) - 0.3040 * (log(*bpd*/10)^2) + 0.9675 * log (*ac*/10) + 0.3557 * log(*arm_vol*))	0.9620	247.8761 ± 206.1607
Lee 5	*weight* = exp.(2.1264 + 1.1461 * log(*ac*/10) + 0.4314 * log(*thigh_vol*))	0.9514	289.2660 ± 234.0763
Lee 2	*weight* = exp.(−3.6138 + 4.6761 * log(*ac*/10) - 0.4959 * (log(*ac*/10)^2) + 0.3795 * log (*arm_vol*))	0.9472	316.4974 ± 242.7964
Ho 6	*weight* = −882.7049 + 73.9955 * *thigh_vol* - 0.497 * (*thigh_vol*^2) + 0.0014 * (*thigh_vol*^3)	0.9385	−7.5001 ± 260.4596
Lee 4	*weight* = exp.(4.7806 + 0.7596 * log(*thigh_vol*))	0.9298	737.4932 ± 344.1904
Lee 1	*weight* = exp.(4.9588 + 1.0721 * log(*arm_vol*) - 0.0526 * (log(*arm_vol*)^2))	0.9281	867.0836 ± 309.5779
Chang	*weight* = 1080.8735 + 22.44701 * *thigh_vol*	0.9229	456.5168 ± 298.2517

Table 7. Comparison of weight estimation with 3D sample.

Formula	Expression	R	Error range
NHD 1 (2D sample)	log10(*weight*) = −3.715073 + 1.873457 * log10(*bpd*) + 0.363783 * log10 (*fl*) + 0.691683 * log10(*ac*) + 0.722245 * log10(*age*)	0.9674	−5.6422 ± 202.0395
NHD 2 (2D sample)	log10(*weight*) = −3.761798 + 2.001731 * log10(*bpd*) + 0.811078 * log10 (*ac*) + 0.826279 * log10(*age*)	0.9667	−5.6111 ± 204.1477
NHD 3 (3D sample)	*weight* = −4988.000528 + 66.374156 * *age* + 0.370084 * *hc* + 1.943247 * *ac* + 39.464816 * *bpd* + 13.215505 * *fl* + 3.658463 * *thigh_vol*	0.9715	0 ± 178.8091
NHD 4 (3D sample)	*weight* = −4982.099978 + 68.089354 * *age* + 2.001675 * *ac* + 39.85375 * *bpd* + 13.229377 * *fl* + 3.619405 * *thigh_vol*	0.9714	0 ± 178.9114
NH 3 (2D sample)	log(*weight*) = −10.047381 + 1.94864 * log(*bpd*) + 0.263745 * log(*hc*) + 0.601972 * log(*fl*) + 0.905524 * log(*ac*)	0.9636	−7.4656 ± 212.5573
NH 4 (2D sample)	log(*weight*) = 3.957543 + 0.02373 * *bpd* + 0.000802 * *hc* + 0.009403 * *fl* + 0.003157 * *ac*	0.9635	−6.0901 ± 214.1153
NH 7 (3D sample)	*weight* = −3617.936175 + 0.513171 * *hc* + 1.960176 * *ac* + 39.804645 * *bpd* + 17.016936 * *fl* + 8.366404 * *thigh_vol* + 5.828808 * *arm_vol*	0.9708	−0.0001 ± 180.9803
NH 8 (3D sample)	*weight* = −3626.314419 + 43.426744 * *bpd* + 23.645338 * *fl* + 11.414273 * *thigh_vol*	0.9698	0 ± 184.0439

Table 8. Weight estimation dual formulas.

Although our formulas are better than all remaining ones with high adequacy (large R) and high accuracy (small error range), other researches are always significant because their formulas are very simple and practical. Moreover, our formulas are not global. If they are applied into other samples collected in other communities, their accuracy may be decreased and they may not be still better than traditional formulas such as Sherpard and Hadlock. However, it is easy to draw from our experimental results that if Phoebe framework is used for the same samples with other researches, it will always produce preeminent formulas. In order to achieve global optimality with Phoebe framework, the following are two essential suggestions:

- Experimenting on Phoebe framework with many samples.
- Adding more knowledge of pregnancy study, ultrasound technique, and obstetrics into Phoebe framework. In other words, the additional knowledge will be modeled as constraints of SG algorithm.

These suggestions go beyond this research. In my opinion, we cannot reach absolutely the global optimality because Phoebe framework focuses on local optimality with specific communities. Essentially, the suggestions only alleviate the weak point of the built-in SG algorithm in global optimality.

6. A proposal of early weight estimation

The used ultrasound samples are collected in fetal age from 28 to 42 weeks because delivery time is not over 48 h since last ultrasound scan. Hence, accuracy of weight estimation is only ensured when ultrasound examinations are performed after 28-week old fetal age. This section proposes an early weight estimation, in which ultrasound measures can be taken before 28-week old fetal age. We do not ensure improvement of estimation accuracy yet because we do not make experiments on the proposal yet, but the gestational sample can be totally collected at any appropriate time points in gestational period. In other words, the sample can lack fetal weights. This is a convenience for practitioners because they do not need to concern fetal weights when taking ultrasound examinations. Consequently, early weight estimation is achieved. As a convention, vectors are column vectors if there is no additional information.

Without loss of generality, regression models are linear such as $Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_n X_n$ and $Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n$ where Y is fetal age and Z is fetal weight, whereas X_i (s) are gestational ultrasound measures such as *bpd*, *hc*, *ac*, and *fl*. Suppose both Y and Z conform normal distribution, according to Eq. (3) ([17] pp. 8–9).

$$P\langle Y|X,\alpha\rangle = \frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left(\frac{-\left(Y-\alpha^T X\right)^2}{2\sigma_1^2}\right)$$
$$P\langle Z|X,\beta\rangle = \frac{1}{\sqrt{2\pi\sigma_2^2}}\exp\left(\frac{-\left(Z-\beta^T X\right)^2}{2\sigma_2^2}\right) \quad (3)$$

where $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_n)^T$ and $\beta = (\beta_0, \beta_1, \ldots, \beta_n)^T$ are parameter vectors where $X = (1, X_1, X_2, \ldots, X_n)^T$ is data vector. The means of Y and Z are $\alpha^T X$ and $\beta^T X$, respectively, whereas the variances

of Y and Z are σ_1^2 and σ_2^2, respectively. Note that the superscript "^{T}" denotes transposition operator in vector and matrix. Let $\boldsymbol{D} = (\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{z})$ be collected sample in which $\boldsymbol{X}$ is a set of sample measures, $\boldsymbol{y}$ is a set of sample fetal ages, and $\boldsymbol{z}$ is a set of fetal weights with note that $\boldsymbol{z}$ is missed (empty) or incomplete. If $\boldsymbol{z}$ is empty, there is no z_i in $\boldsymbol{z}$. If $\boldsymbol{z}$ is incomplete, $\boldsymbol{z}$ has some values but there are also some missing values in $\boldsymbol{z}$. However, the constraint is that $\boldsymbol{y}$ must be complete, which means that all pregnant women within the research knew their gestational age. Now we focus on estimate α and β based on $\boldsymbol{D}$. As a convention, let α^* and β^* be estimates of α and β, respectively ([17] p. 8).

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \vdots \\ \boldsymbol{x}_N^T \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1n} \\ 1 & x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & x_{N2} & \cdots & x_{Nn} \end{pmatrix}$$

$$\boldsymbol{x}_i = \begin{pmatrix} 1 \\ x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}, \boldsymbol{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{pmatrix}$$

Given X, joint probability of Y and Z is product of the probability of Y given X and the probability of Z given X because Y and Z are conditionally independent given X, according to Eq. (4).

$$P\langle Y, Z|X, \alpha, \beta\rangle = P\langle Y|X, \alpha\rangle P\langle Z|X, \beta\rangle = \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2}} \exp\left(\frac{-\left(Y - \alpha^T X\right)^2}{2\sigma_1^2} - \frac{\left(Z - \beta^T X\right)^2}{2\sigma_2^2}\right) \quad (4)$$

Conditional expectation of sufficient statistic Z given X with regard to $P(Z \mid X, \beta)$ is specified by Eq. (5).

$$E\langle Z|X\rangle = \beta^T X \quad (5)$$

When Z is hidden variable, there is a latent dependent relationship between Y and Z, which is specified by joint probability of Y and Z.

$$P(Y, Z) = P(Y)P\langle Z|Y\rangle$$

Variables Y and Z have different measures. For instance, the unit of Y is week, whereas the unit of Z is gram. Suppose Y is considered as discrete variable whose values from 1 to K where K can be up to 42, for example. The $P(Y)$ becomes parameter θ_Y, which is the probability of Y where Y is from 1 to K.

$$P(Y, Z) = \theta_Y P\langle Z|Y\rangle$$

For each Z, suppose the condition probability $P(Z \mid Y)$ is distributed normally with mean μ_Y and variance ${\sigma_Y}^2$. Eq. (6) specifies the joint probability $P(Y, Z)$.

$$P\langle Y, Z|\theta_Y, \mu_Y, \sigma_Y^2\rangle = \frac{\theta_Y}{\sqrt{2\pi\sigma_Y^2}}\exp\left(\frac{-(Z-\mu_Y)^2}{2\sigma_Y^2}\right) \tag{6}$$

Conditional expectation of sufficient statistic Z given Y with regard to $P(Z \mid Y, \mu_Y, {\sigma_Y}^2)$ is specified by Eq. (7).

$$E\langle Z|Y\rangle = \mu_Y \tag{7}$$

Please pay attention to Eq. (7) because Z will be estimated by such expectation later. Eq. (8) specifies expectation of sufficient statistic Z with regard to $P(Y, Z \mid \theta_Y, \mu_Y, {\sigma_Y}^2)$.

$$E(Z) = \sum_{Y=1}^{K}\theta_Y\mu_Y \tag{8}$$

Due to:

$$E\langle Z|\theta_Y, \mu_Y, \sigma_Y^2\rangle = \sum_{Y=1}^{K}\int_Z ZP\langle Y, Z|\theta_Y, \mu_Y, \sigma_Y^2\rangle dZ = \sum_{Y=1}^{K}\theta_Y E\langle Z|\theta_Y, \mu_Y, \sigma_Y^2\rangle = \sum_{Y=1}^{K}\theta_Y\mu_Y$$

The full joint probability of Y and Z given X and parameters α, β, θ_Y, μ_Y, and ${\sigma_Y}^2$ is the product specified by Eq. (9).

$$\begin{aligned} &P\langle Y, Z|X, \alpha, \beta, \theta_Y, \mu_Y, \sigma_Y^2\rangle = P\langle Y, Z|\theta_Y, \mu_Y, \sigma_Y^2\rangle P\langle Y, Z|X, \alpha, \beta\rangle \\ &= P\langle Y, Z|\theta_Y, \mu_Y, \sigma_Y^2\rangle P\langle Y|X, \alpha\rangle P\langle Z|X, \beta\rangle \end{aligned} \tag{9}$$

where $P(Y, Z \mid X, \alpha, \beta)$ and $P(Y, Z \mid \theta_Y, \mu_Y, {\sigma_Y}^2)$ are specified by Eqs. (4) and (6), respectively. Eq. (9) indicates that both explicit dependence via $P(Y, Z \mid X, \alpha, \beta)$ and implicit dependence via $P(Y, Z \mid \theta_Y, \mu_Y, {\sigma_Y}^2)$ between Y and Z. Explicit dependence and implicit dependence share equal influence on Z if $E(Z \mid X)$ specified by Eq. (5) is equal to $E(Z)$ specified by Eq. (8), according to Eq. (10).

$$\sum_{Y=1}^{K}\theta_Y\mu_Y = \beta^T X \tag{10}$$

Given sample $\boldsymbol{D}$, all θ_Y become constants and determined by Eq. (11).

$$\theta_Y = \frac{\text{The number of } y_i = Y}{N} \tag{11}$$

For convenience, let $\Theta = (\alpha, \beta, \mu_Y)^T$ be the compound parameter. The full joint probability specified by Eq. (9) is rewritten as follows:

$$P\langle y, z|X, \Theta\rangle = P\langle y, z|\mu_Y, \sigma_Y^2\rangle P\langle y|X, \alpha\rangle P\langle z|X, \beta\rangle$$

$$= \prod_{i=1}^{N} P\langle y_i, z_i|\mu_Y, \sigma_Y^2\rangle P\langle y_i|\boldsymbol{x_i}, \alpha\rangle P\langle z_i|\boldsymbol{x_i}, \beta\rangle$$

(Due to all observations are independently and identically distributed)

$$= \left(\frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2}}\right)^N * \exp\left(\frac{-1}{2}\left(\sum_{i=1}^{N}\frac{\left(y_i - \alpha^T\boldsymbol{x_i}\right)^2}{\sigma_1^2} + \sum_{i=1}^{N}\frac{\left(z_i - \beta^T\boldsymbol{x_i}\right)^2}{\sigma_2^2}\right)\right) * \prod_{i=1}^{N}\prod_{Y=1}^{K}\frac{\delta(y_i, Y)\theta_Y}{\sqrt{2\pi\sigma_Y^2}}\exp\left(\frac{-\left(z_i - \mu_Y\right)^2}{2\sigma_Y^2}\right)$$

where

$$\delta(y_i, Y) = \left\{\begin{array}{l} 1 \text{ if } y_i = Y \\ 0 \text{ if } y_i = Y \end{array}\right\}$$

It is conventional that if $\delta(y_i, Y) = 0$ then, the respective probability $P(y_i, z_i \mid \mu_Y, {\sigma_Y}^2)$ is removed from the product. The log-likelihood function is logarithm of the full joint probability as follows:

$$L(\Theta) = \log(P\langle y, z|X, \Theta\rangle) = -N\log(2\pi) - \frac{N\log\left(\sigma_1^2\right)}{2} - \frac{N\log\left(\sigma_2^2\right)}{2}$$

$$-\frac{1}{2\sigma_1^2}\sum_{i=1}^{N}\left(y_i - \alpha^T\boldsymbol{x_i}\right)^2 - \frac{1}{2\sigma_2^2}\sum_{i=1}^{N}\left(z_i - \beta^T\boldsymbol{x_i}\right)^2$$

$$+\sum_{i=1}^{N}\sum_{Y=1}^{K}\delta\left(y_i, Y\right)\left(\log(\theta_Y) - \frac{\log(2\pi)}{2} - \frac{\log\left(\sigma_Y^2\right)}{2} - \frac{\left(z_i - \mu_Y\right)^2}{2\sigma_Y^2}\right)$$

When $\log(2\pi)$ and θ_Y are constants, the reduced log-likelihood function is derived from the log-likelihood as seen in Eq. (12).

$$\begin{aligned} l(\Theta) = &-\frac{N}{2}\log\left(\sigma_1^2\right) - \frac{N}{2}\log\left(\sigma_2^2\right) - \frac{1}{2\sigma_1^2}\sum_{i=1}^{N}\left(y_i - \alpha^T\boldsymbol{x_i}\right)^2 - \frac{1}{2\sigma_2^2}\sum_{i=1}^{N}\left(z_i - \beta^T\boldsymbol{x_i}\right)^2 \\ &-\frac{1}{2}\sum_{i=1}^{N}\sum_{Y=1}^{K}\delta\left(y_i, Y\right)\left(\log\left(\sigma_Y^2\right) + \frac{\left(z_i - \mu_Y\right)^2}{\sigma_Y^2}\right) \end{aligned} \tag{12}$$

The optimal estimate Θ^* is a maximizer of $l(\Theta)$, according to Eq. (13) ([17] p. 9).

$$\Theta^* = \underset{\Theta}{\operatorname{argmax}}\, L(\Theta) = \underset{\Theta}{\operatorname{argmax}}\, l(\Theta) \tag{13}$$

By taking first-order partial derivatives of $l(\Theta)$ with regard to Θ ([18] p. 34), we obtain:

$$\frac{\partial l(\Theta)}{\partial \alpha} = \frac{1}{\sigma_1^2}\sum_{i=1}^{N}\left(y_i - \alpha^T \boldsymbol{x_i}\right)(\boldsymbol{x_i})^T$$

$$\frac{\partial l(\Theta)}{\partial \beta} = \frac{1}{\sigma_2^2}\sum_{i=1}^{N}\left(z_i - \beta^T \boldsymbol{x_i}\right)(\boldsymbol{x_i})^T$$

$$\frac{\partial l(\Theta)}{\partial \mu_Y} = \sigma_Y^2\sum_{i=1}^{N}\delta\left(y_i, Y\right)\left(z_i - \mu_Y\right)$$

When first-order partial derivatives of $l(\Theta)$ are equal to zero, it gets locally maximal. In other words, Θ^* is solution of the equation system 14 resulted from setting such derivatives to be zero and setting $E(Z \mid X) = E(Z)$.

$$\left\{\begin{array}{l} \sum_{i=1}^{N}\left(y_i - \alpha^T \boldsymbol{x_i}\right)(\boldsymbol{x_i})^T = \boldsymbol{0}^T \\ \sum_{i=1}^{N}\left(z_i - \beta^T \boldsymbol{x_i}\right)(\boldsymbol{x_i})^T = \boldsymbol{0}^T \\ \sum_{i=1}^{N}\delta\left(y_i, Y\right)\left(z_i - \mu_Y\right) = 0 \\ \sum_{j=1}^{K}\theta_j\mu_j = \beta^T \boldsymbol{x_i} \text{ for some } i \end{array}\right\} \tag{14}$$

where

$$\delta\left(y_i, Y\right) = \left\{\begin{array}{l} 1 \text{ if } y_i = Y \\ 0 \text{ if } y_i \neq Y \end{array}\right\}$$

The notation $\boldsymbol{0} = (0, 0, \ldots, 0)^T$ denotes zero vector. All equations in system 14 are linear, whose unknowns are $\Theta = (\alpha, \beta, \mu_Y)^T$. The last equation in system 14 is Eq. (10) with the heuristic assumption that explicit dependence and implicit dependence share equal influence on Z. Such last equation is only used to adjust μ_Y (s) if the heuristic assumption is concerned; otherwise it is ignored.

We apply expectation maximization (EM) algorithm into estimating $\Theta = (\alpha, \beta, \mu_Y)^T$ with lack of fetal weights. Note that the full joint probability $P(Y, Z \mid X, \alpha, \beta, \mu_Y)$ specified by Eq. (9) is product of regular exponential distributions. EM algorithm has many iterations, and each iteration has expectation step (E-step) and maximization step (M-step) for estimating parameters. Given current parameter $\Theta^t = (\alpha^t, \beta^t, \mu_Y{}^t)^T$ at the t^{th} iteration, the two steps are shown in **Table 9** ([19] p. 4).

The equation system 14 is solvable because missing values z_i (s) were estimated in E-step. The EM algorithm stops if at some t^{th} iteration, we have $\Theta^t = \Theta^{t+1} = \Theta^*$. At that time, $\Theta^* = (\alpha^*, \beta^*, \mu_Y{}^*)^T$ is the optimal estimate of EM algorithm, and hence, linear regression functions of Y and Z are determined with α^*, β^*.

1.	E-step: Estimating only missing values z_i (s) as the expectation of themselves based on the current mean $\mu_{y_i}^t$, according to Eq. (7). Note, each missing value z_i is always associated with an observation y_i. $$z_i = E\langle z_i \mid y_i \rangle = \mu_{y_i}^t$$
2.	M-step: The next parameter Θ^{t+1} is a maximizer of $l(\Theta)$, which is the solution of equation system 14. Note, Θ^{t+1} becomes current parameter for the next iteration.

Table 9. E-step and M-step of EM algorithm.

As usual, all parameters are changed after every iteration of EM algorithm, but fortunately, α^* is determined as a partial solution of equation system 14 at the first iteration of EM process because both $\boldsymbol{X}$ and $\boldsymbol{y}$ are complete. In other words, α^* is fixed, whereas β and μ_Y are changed in EM process. Eq. (15) ([20] p. 417) specifies α^*.

$$\alpha^* = \alpha^1 = \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y} \tag{15}$$

where the superscript "$^{-1}$" denotes the inversion of matrix.

At the first iteration, as usual Θ^1 is initialized arbitrarily but we can improve convergence of EM algorithm by initializing μ_Y^1 as sample mean. Without loss of generality, suppose practitioners obtained $n < N$ fetal weights $z_1, z_2, \ldots, z_n$ from n ultrasound scans. Moreover, the fetal age of all pregnant women over such n scans is the same, which is Y. Thus, μ_Y^1 is initialized by Eq. (16).

$$\mu_Y^1 = \frac{1}{n}\sum_{i=1}^{n} z_i \tag{16}$$

The parameter β^1 at the first iteration is initialized according to previous studies in the literature.

7. Conclusions

According to experimental results, there is no doubt that Phoebe framework produces optimal formulas with high adequacy and accuracy; please see **Tables 4–8** for more details. However, we also recognize the weak point of our research is that the built-in SG algorithm can lose some good formulas due to the heuristic conditions. The suggestive solution is to add more constraints in such conditions; please read the article "A framework of fetal age and weight estimation" ([10] pp. 24–25) for more details. The proposal of early weight estimation uses actually an additional constraint which is the latent relationship between fetal age and fetal weight. Such latent relationship represented by the joint probability of fetal age and weight is a knowledge aspect of pregnancy study. For further research, we will make experiment on the proposal and try our best to discover other knowledge aspects.

Another weak point of our research is difficult to apply our complex formulas for fast mental calculation because we must pay the price for their high accuracy. In the future, we will embed

these formulas into software or hardware of medical ultrasound machine so that users are easy to read estimated values resulted from machine.

Acknowledgements

We express our deep gratitude to the author Michael Thomas Flanagan – University College London and the author Jos de Jong for giving us helpful software packages that help us to implement the framework.

Author details

Loc Nguyen[1]*, Truong-Duyet Phan[2] and Thu-Hang T. Ho[3]

*Address all correspondence to: ng_phloc@yahoo.com

1 Sunflower Soft Company, Ho Chi Minh, Vietnam

2 Hanoi Medical University, Hanoi, Vietnam

3 Vinh Long General Hospital, Vinh Long, Vietnam

References

[1] Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK. Estimation of fetal weight with use of head, body and femur measurements: A prospective study. American Journal of Obstetrics and Gynecology. 1 February 1985;**151**(3):pp. 333-337

[2] Phan DT. Ứng dụng siêu âm để chẩn đoán tuổi thai và cân nặng thai trong tử cung. Hanoi: Hanoi University of Medicine; 1985

[3] Phạm TNT. Ước lượng cân nặng thai nhi qua các số đo của thai trên siêu âm. Ho Chi Minh: Ho Chi Minh University of Medicine and Pharmacy; 2000

[4] Ho THT. Nghiên Cứu Phương Pháp Ước Lượng Trọng Lượng Thai, Tuổi Thai Bằng Siêu Âm Hai và Ba Chiều. Hanoi: Hanoi Univerisy of Medicine; 2011

[5] Shepard JM, Richards AV, Berkowitz LR, Warsof LS, Hobbins CJ. An evaluation of two equations for predicting fetal weight by ultrasound. American Journal of Obstetrics and Gynecology. 1 January 1982;**142**(1):47-54

[6] Campbell S, Wilkin D. Ultrasonic measurement of fetal abdomen circumference in the estimation of fetal weight. BJOG: An International Journal of Obstetrics & Gynecology. September 1975;**82**(9):689-697

[7] Lee W, Balasubramaniam M, Deter RL, Yeo L, Hassan SS, Gotsch F, Kusanovic JP, Gonçalves LF, Romero R. New fetal weight estimation models using fractional limb volume. Ultrasound in Obstetrics & Gynecology. 1 November 2009;**34**(5):556-565

[8] Chang F-M, Liang R-I, Ko H-C, Yao B-L, Chang C-H, Yu C-H. Three-dimensional ultrasound-assessed fetal thigh volumetry in predicting birth weight. Obstetrics & Gynecology. September 1997;**90**(3):331-339

[9] Varol F, Saltik A, Kaplan PB, Kilic T, Yardim T. Evaluation of gestational age based on ultrasound fetal growth measurements. Yonsei Medical Journal. June 2001;**42**(3):299-303

[10] Flanagan MT. In: Flanagan MT, editor. Java Scientific Library. London, England: University College London; 2004

[11] Jong Jd, A Java Expression Parser, Rotterdam: SpeQ Mathematics; 2010

[12] Oracle, "Java language," Oracle Corporation, [Online]. Available: https://www.oracle.com/java. [Accessed 25 December 2014]

[13] Nguyen L, Ho H. A framework of fetal age and weight estimation. Journal of Gynecology and Obstetrics (JGO). 30 March 2014;**2**(2):pp. 20-25

[14] Ho THT, Phan DT. Ước lượng cân nặng của thai từ 37–42 tuần bằng siêu âm 2 chiều. Journal of Practical Medicine. December 2011;**12**(797):8-9

[15] Ho T-HT, Phan DT. Ước lượng tuổi thai qua các số đo thể tích cánh tay bằng siêu âm 3 chiều và các số đo bằng siêu âm 2 chiều. Journal of Practical Medicine. December 2011; **12**(798):12-15

[16] Nguyen L, Ho T-HT. Experimental results of phoebe framework: optimal formulas for estimating fetus weight and age. Journal of Community & Public Health Nursing13 March 2017;**3**(2):1-5

[17] Lindsten F, Schön TB, Svensson A, Wahlström N. Probabilistic Modeling – Linear Regression & Gaussian processes. Uppsala: Uppsala University; 2017

[18] Nguyen L. In: Evans C, editor. Matrix Analysis and Calculus. 1st ed. Hanoi: Lambert Academic Publishing; 2015. p. 72

[19] Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B (Methodological). 1977; **39**(1):1-38

[20] Montgomery DC, Runger GC. Applied Statistics and Probability for Engineers. 3rd ed. New York, NY: John Wiley & Sons, Inc.; 2003. p. 706

Chapter 8

Using Patient Registries to Identify Triggers of Rare Diseases

Feras M. Ghazawi, Steven J. Glassman,
Denis Sasseville and Ivan V. Litvinov

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.76449

Abstract

Mapping the distribution of patients and analyzing disease clusters is an effective method in epidemiology, where the non-random aggregation of patients is carefully investigated. This can aid in the search for clues to the etiology of diseases, particularly the rare ones. Indeed, with the increased incidence of rare diseases in certain populations and/or geographic areas and with proper analysis of common exposures, it is possible to identify the likely promoters/triggers of these diseases at a given time. In this chapter, we will highlight the appropriate methodology and demonstrate several examples of cluster analyses that lead to the recognition of environmental, occupational and communicable preventable triggers of several rare diseases.

Keywords: cluster investigation, epidemiology, rare diseases, exposures, disease triggers, patient registries

1. Introduction

Many diseases are preventable with lifestyle modifications and by minimizing exposures to harmful substances. In fact, it was recently reported that nearly half of all cancer-related deaths in the United States were attributable to modifiable and preventable risk factors [1]. Through epidemiological studies and careful examination of public health data such as disease registries, and by studying disease distribution, incidence, prevalence and mortality trends, the occurrence of diseases in defined populations can be estimated and can be related to different external factors. Disease clusters are aggregates of patients with a particular disease in a specified time period and at a defined geographical level, occurring at a rate markedly higher than

expected. Analyzing and mapping the incidence rates of diseases indeed can help identify non-random distributions of patient clusters, while a proper assessment of the population demographics as well as the surrounding environment can implicate occupational, communicable and environmental exposures as potential causes for a given disease. For instance, in the 1800's, despite limited knowledge on the etiology of many diseases such as cholera, clustering analysis was the method that enabled physicians and scientists to establish a definite link between the disease outbreaks and causative or potentiating agents from the surrounding environment. In the case of cholera, water from a contaminated pump infected with *Vibrio cholerae* bacterium was clearly identified as a disease source in London, England. As highlighted in the examples in Section 2.2, it will be evident to the reader that geographic clustering analyses of patient populations can shed light on the triggers of many rare diseases.

2. Cluster investigation analysis

2.1. Cluster investigations

In epidemiology, trends and causes of diseases and their progression and regression rates can be monitored over time and the occurrence of diseases within the defined populations can be estimated. There are several types of epidemiologic studies including cohort, case–control, cross-sectional, ecologic and cluster studies. Epidemiological studies have a significant impact on public health outcomes as they identify increased disease incidence/prevalence rates, which shape health policies including preventative measures and resource allocation planning, accordingly. Spatial epidemiology is the description and geographical analysis of health data, taking into account patients' demographics, and risk factors including socioeconomic, genetic, environmental, behavioral, infectious and noninfectious exposures [2]. Detection of disease clusters is an integral component of spatial epidemiology as it identifies disproportionately high rates of a disease in a given population, which ultimately generates hypotheses that can help elucidate disease triggers/promoters.

Clustering analyses can be characterized as either general (non-focused) or specific (focused). In a general clustering analysis, the precise location of disease clusters is not studied but rather the clustering tendency of the disease and the overall distribution of disease is examined [3, 4]. On the other hand, specific disease clustering analysis carefully describes unusual nonrandom accumulation of disease outbreaks and the precise location of clusters, in time or space, that are unlikely to be due to chance alone [3, 4]. These investigations can be applied to formulate hypotheses and elucidate potential causes of diseases. Further, clustering patterns of diseases have many applications beyond identifying disease triggers, including identification of areas with high disease prevalence in order to optimize medical management and resource allocation. This will be discussed in Section 2.2.

2.2. Applications of cluster investigations in identifying disease triggers

2.2.1. Cholera

Cholera is an acute infectious diarrheal disease that can be fatal within days if left untreated. This disease became a major health threat in the 1800s, with several outbreaks that had

devastating outcomes. The accepted explanation of cholera outbreaks at the time was attributed to the "Miasma theory", which suggested that poisonous vapor or mist filled with substances from decomposed matter ("Miasmata") caused many diseases including cholera, chlamydia and plague [5]. In 1854, a severe outbreak of cholera occurred in London, England, killing more than 600 people. Dr. John Snow, an English physician, investigated the cause of this epidemic by analyzing the geographic distribution of cholera and plotting cholera cases on a map, along with certain landmarks in the city including providers of potable water (**Figure 1**). Notably, most of the cholera cases occurred within 250 yards of the intersection of Broad and Cambridge streets and in close proximity to a public water pump on Broad street. This observation prompted the local council to disable the water pump which halted the spread of cholera. This analysis enabled the identification of the precise source of the cholera outbreak in London as the public water pump, which was built near an old open toilet. This work established for the first time that cholera can spread via contaminated water [6]. The breakthrough in fact paved the way for the field of epidemiology.

2.2.2. Mesothelioma

Mesothelioma is a rare but aggressive cancer that arises in the mesothelium, the lining of the pleura, peritoneum, and pericardium. Studying the prevalence of mesotheliomas in asbestos miners of South Africa established asbestos exposure as a critical factor responsible for this deadly malignancy. In a study by Wagner and colleagues, it was noted that while mesothelioma is a very rare disease in the Northwest Cape province of South Africa, 33 cases were described in the area, each having occupational exposure to crocidolite asbestos mining [8]. This finding was shortly followed by several population studies in Quebec Canada, United Kingdom, The Netherlands, Germany, Scotland and Northern Ireland, demonstrating that

Figure 1. Original map by Dr. John Snow illustrating clustering of cholera cases in the London epidemic of 1854. Cholera cases are highlighted as black lines [7].

most of the described mesothelioma patients clustered in several communities where occupational exposure to asbestos was routine. At that time asbestos was commonly used in insulation, construction, factory work as well as in shipyards. This analysis confirmed the causal link between asbestos and mesotheliomas and led to a legislative action to ban the use of this carcinogen in construction and other workplaces [9].

2.2.3. Squamous cell carcinoma

For many centuries, arsenic was used by the Egyptians, Greeks, Asians and Romans for many applications including the treatment of rheumatism and for facial hair removal. Little was known about the carcinogenic effects of arsenic at that time. In 1898, Geyer conducted a detailed population study in Reichenstein, Silesia (Prussia). In this small arsenic-mining town, chronic poisoning took place primarily through the use of drinking water contaminated by precipitating arsenical fumes in the rain. This work shed light on the carcinogenic effects of arsenic [10]. Affected individuals developed a constellation of symptoms including pigmentation changes and hyperkeratosis (wart-like lesions) on the palms and soles. The latter had a high risk of progression to cutaneous squamous cell carcinomas. This condition was referred to as "Reichenstein's disease" [11]. The significant increase of this disease in the residents of this town helped establish the link between arsenic and the occurrence of arsenical keratoses and squamous cell carcinomas of the skin. This example further demonstrates the importance of non-random clustering of rare diseases in identifying novel environmental or occupational disease triggers.

2.2.4. Cutaneous T-cell lymphoma

Cutaneous T-cell lymphoma (CTCL) is a rare group of non-Hodgkin lymphomas that primarily involves the skin. Patients with CTCL typically present with persistent, red itchy patches and thickened plaques that are located mostly on the trunk. As the malignancy progresses, patients can develop skin tumors with concomitant involvement of lymph nodes and visceral organs. In some stages, the disease involves the blood and patients can develop erythroderma (generalized redness and desquamation of the skin) and suffer intractable pruritus as well as B-symptoms of lymphoma. Many advanced disease patients succumb to this malignancy within 2–3 years. Unfortunately, the risk factors and promoters for this disease remained poorly understood for many years. It is recognized that disruption of molecular pathways in skin lymphocytes by bacterial, viral or environmental factors can lead to cutaneous lymphomas [12–14]. Although progress has been made in the past few decades, the precise pathogenesis by which CTCL develops remains poorly understood. Several reports from different parts of the world examined the distribution of CTCL patients illustrating non-random clustering of cases. This was shown in Sweden [15], Houston, Texas (**Figure 2**) [16, 17] and the Pittsburgh metropolitan area [18]. Furthermore, the unusually high incidence of CTCL in married couples [19], and in families [20] was also noted. These clustering patterns of CTCL patients strongly argue for the existence of external and potentially preventable risk factors for this rare skin cancer.

Several factors have been implicated in CTCL carcinogenesis, including immunosuppression, vitamin D deficiency, bacterial agents (*Staphylococcus aureus*, *Mycobacterium leprae* and *Chlamydophila pneumoniae*), medications (calcium channel blockers, angiotensin converting

enzyme inhibitors, hydrochlorothiazide, and serotonin reuptake inhibitors), dermatophytes and viruses (EBV, HSV and HTLV-1) [21]. However, none of these agents have been definitively linked with this skin lymphoma.

In addition, recent studies in Canada further confirmed the existence of disease clusters and areas completely spared by this malignancy and implicated industrial exposure and living in a proximity to major transportation junctions as potential triggers for CTCL [22, 23]. Considering that the majority of skin cancers are caused by external and often preventable triggers (e.g. UV radiation, HPV, polyomaviruses, etc.) it is not surprising that skin lymphomas

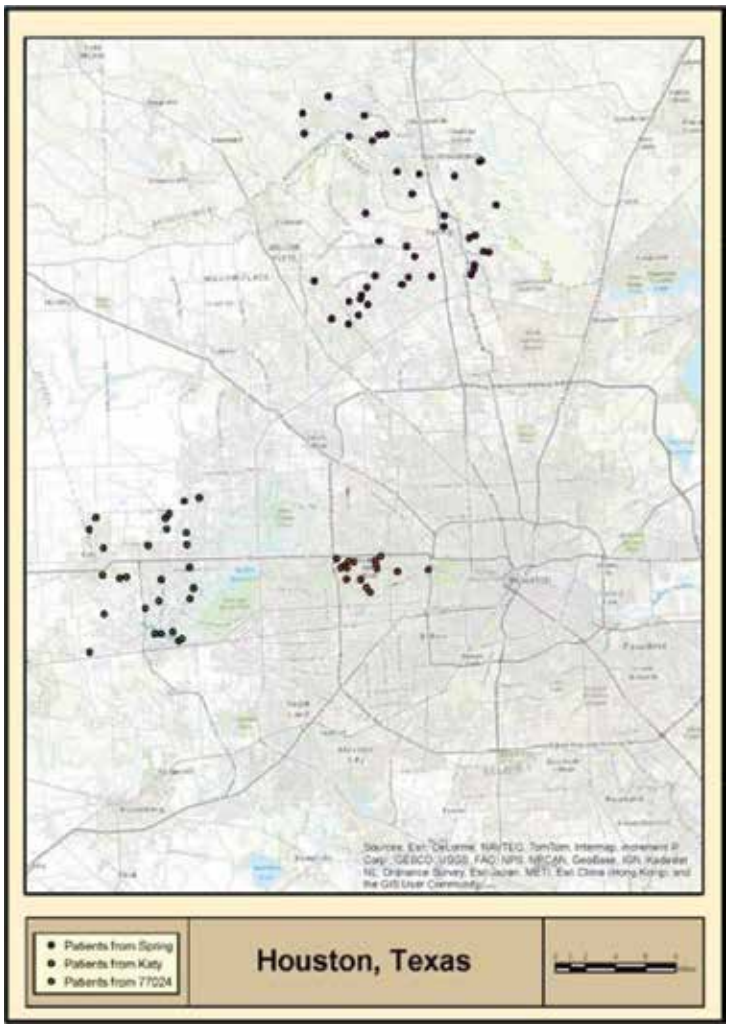

Figure 2. Geographic mapping of cutaneous T-cell lymphoma cases in Houston, Texas demonstrating clustering pattern of patients with this rare malignancy with incidence rates 5–20-fold higher than expected. Patients in 'Spring' community are indicated by violet, patients in 'Katy' community are indicated by green, and patients residing in the Houston memorial area are indicated by orange. Adapted from [17].

could also be caused by an external trigger. Currently, the search for such trigger(s) for this malignancy is ongoing.

2.2.5. Childhood leukemia

Another example revealing a cause of an important disease came from an observation in the early 1980s in Woburn, Massachusetts, where an elevated incidence rate of childhood leukemia was documented. An extensive investigation of the geographical distribution of these patients helped implicate chlorinated organic compounds contaminating two of eight municipal wells servicing Woburn as a cause of childhood leukemia. Specifically, it was shown that select dwellings where the patients with this cancer resided, were provided water from these contaminated wells [24].

2.2.6. Bladder cancer

Bladder cancer is a disease of significant morbidity and mortality [25]. Cluster investigation recently helped identify occupational and behavioral promoters for this cancer. These factors are potentially modifiable and thus rates of this malignancy could possibly be reduced with primary prevention. The astute observation in 1895 by Rehn, a German physician, showed that the incidence rates of bladder cancer were remarkably high in aniline dye industry workers. This was the first evidence that occupational risk factors can be directly implicated in this malignancy [26]. By carefully analyzing the incidence of bladder cancers in industrial workers, it was possible to identify aromatic amines, polycyclic aromatic hydrocarbons and chlorinated hydrocarbons that are now well recognized as causative agents for this disease [25].

2.2.7. Emerging trends

2.2.7.1. Multiple sclerosis

Multiple sclerosis (MS) is an autoimmune demyelinating disease, affecting the central nervous system and resulting in a spectrum of neurological symptoms including vision problems, fatigue, pains, spasms and cognitive decline. The precise triggers of this rare disease have not yet been described or identified. However, studying the epidemiology and geographic distribution of MS globally has yielded many interesting trends that allowed generation of a number of hypotheses addressing the cause of MS. Clusters of new MS cases have been reported in many communities around the world including the United States, Canada, Europe, Israel, New Zealand, Australia and Russia [27–30]. Many studies indicated significant variation in the global distribution of MS patients, where the incidence of this autoimmune disease is relatively uncommon in tropical climates, but is much more common in temperate zones and in the Western Hemisphere [31]. Furthermore, remarkably elevated incidence rates in northern latitudes were reported [32, 33]. Many theories have been postulated to implicate promoters of MS, such as diet, soil minerals and deficiency in vitamin D [32, 33]. The identity of a definite trigger for MS remains unknown, and extensive follow up of identified clusters may potentially provide some clues in the future.

2.2.7.2. Alzheimer's disease

Alzheimer's disease is a common, yet incompletely understood form of dementia. Differences in the geographical distribution of patients with Alzheimer's disease were reported, highlighting the possible contribution of nutritional or socio-environmental factors in the development and progression of the disease [34]. Indeed, levels of essential trace elements including selenium, magnesium, iron, copper and zinc were shown to be markedly reduced in Alzheimer's patients compared to same age healthy individuals [35]. This illustrates that further epidemiologic studies can be used to associate nutritional deficiencies with diseases.

2.3. Applications of cluster investigations in identifying nutritional deficiencies

2.3.1. Scurvy

Deficiency in micronutrients and vitamins can result in a variety of diseases. For instance, vitamin A deficiency is a known cause of keratomalacia, while vitamin D deficiency in childhood invariably causes rickets. One important use of clustering analysis in epidemiology is to identify nutritional deficiencies.

During the Age of Discovery in the fifteenth and sixteenth century, particularly during long transatlantic journeys, it was noted that the incidence of scurvy, a rare disease caused by a severe deficiency of vitamin C (ascorbic acid), was much higher in sailors, pirates and other sea explorers. Also, the disease later affected soldiers in world wars. Scurvy is characterized by general weakness, gingivitis and bleeding disorders. It was noted that eating citrus fruits prevented and cured this disease in sailors, which enabled later confirmation that vitamin C deficiency is the sole cause of scurvy. Thus, careful demographic and epidemiologic analyses of these individuals, who did not have access to fresh fruit and vegetables, established a link between nutritional deficiency and disease.

2.3.2. Goiter

Thyroid goiters, which represent enlargement of the thyroid gland, are caused by iodine deficiency. Fortification of table salt, medications and common foods like bread with iodine has largely eliminated the once pandemic goiter, but the condition persists in some regions of the developing world. The first hypothesis linking iodine with the treatment of goiter was made in the mid-1800s by a French chemist, Adolphe Chatin [36]. However, fortification of table salt with iodine was not implemented in the United States until the early 1920s [37], and this was, at least in part, driven by epidemiological research.

It was noted that the prevalence of goiter was very high (in approximately 26–70% of children) in the upper Midwest and Great Lakes regions of the United States. In fact, this endemic region was known at the time as the "Goiter Belt" [38]. The prevalence was also reported as high as 64.4% in some areas of Michigan [39]. This highlighted the severity of the problem, sparking a major public health initiative to supplement table salt with iodine. The intervention was very successful, as the incidence of goiter in Michigan dropped by up to 90% within a decade of iodine supplementation [40]. Currently, several areas have remarkably

high prevalence of goiter, such as parts of India and the Himalayan/sub-Himalayan belts [41]. In fact, despite efforts to implement iodine supplementation and table salt fortification with iodine, the goiter prevalence in these communities has not decreased significantly [42]. Thus, more work needs to be done to address logistic, cultural and other obstacles to eliminate suffering from goiter in these regions. In conclusion, recognizing the high prevalence of 'uncommon' diseases such as goiter has important clinical implications. These studies help detect regions with micronutrient deficiency, which can serve as surrogate markers for poor nutrition and encourage prioritizing resource allocation to the affected communities.

2.4. Conducting a proper cluster investigation analysis

2.4.1. Systematic approach to conducting a cluster analysis

The study of the incidence/prevalence of a disease and mapping its distribution requires a systematic approach when trying to implicate occupational and environmental exposures as disease triggers/promoters. Mapping and exposure investigations are critical to highlight the existence and significance of identified clusters. However, it is not enough to only learn about the geographical disease clusters (i.e., disease hot-spots). It is also important to identify regions that are significantly spared by the disease (i.e., disease cold-spots). Detailed epidemiological and statistical analysis of both can help rule-in or rule-out environmental contamination or exposures as disease triggers [43]. A point-by-point guide of a systematic approach to conducting a cluster investigation is provided below:

1. Define the disease and population(s) to be examined.
2. Obtain 'background' information about patient demographics to enable standardization of incidence and mortality rates (such as standardization by age, gender, race, socioeconomic status, etc.).
3. Obtain census or other population information to enable calculating incidence and mortality rates per country, territory/state/province, city and postal code. It is also helpful to learn about common exposures or diseases in that population to adjust for potential confounders. For instance, when studying the incidence of hepatitis C infection in a population, the rate of HIV prevalence would be an important confounder, since in many patients there is co-infection with both viruses due to shared risk factors for viral transmission. Population demographic parameters often vary and can be useful for subsequent analysis of collected data. The specific parameters of interest will differ for each disease, but often include population size, age ranges, race, gender distribution, socioeconomic status, data on lifestyle/behaviors, other environmental, occupational, or local rates of communicable diseases, etc.
4. Obtain public health data on patients with the disease of interest (e.g. local or national cancer registries and Centers for Disease Control, etc.). It is critical to obtain the data from population-based registries since it is often very difficult to draw conclusions from data based on a single medical center or a few select hospitals' experience. One must always seek to correlate single center evidence with population-based registries/databases.

Relevant collected information should include age at diagnosis, year of diagnosis (for incidence calculation), gender, ethnic background (to study disease ethnic predilection), patients' addresses (for geographical mapping), age at death, year of death (for mortality calculation), disease stage, etc.

5. Subsequent calculations of incidence can be easily performed using the obtained data (incidence rate per year = number of new patients per year/population at risk per year). A plot of incidence rates (y axis) *vs.* year (x axis) will enable calculating an average incidence rate and trending the change of rate over time. Mortality calculations are done similarly, using number of deceased patients per year/population at risk.

6. Incidence rates in smaller geographical regions can be calculated similarly. For rare diseases, it is important to include only locations with at least >5000–10,000 residents per geographical area to reduce erroneous false-positive hits, in which a few cases of disease occurring within a scarcely populated area (e.g., <5000 residents) may artificially inflate the incidence/mortality rate.

7. The calculated incidence/mortality rates can be normalized to several variables (such as age, gender, ethnicity) or to a known distribution of relevant disease-specific variables (such as communicable diseases, geographical latitude, socioeconomic status, etc.) This is important to account for potential confounding variables and to highlight trends that can be 'masked' if rates are not normalized in subsequent analyses.

8. Conduct proper statistical analysis to determine statistically significant high and low incidence/mortality rates per geographical region at all levels. Two of the most commonly used methods of statistical analysis are the chi-square test (comparing observed number of cases to that expected under an assumed Poisson distribution) and the Knox test for time–space interaction, among more than 70 different methods, which have been used in previously published studies [43].

9. Plot the incidence rates in a specialized computer program such as ArcGIS or other geographic information system (GIS) software. Generate several maps, choosing appropriate color schemes representing standardized rates. It may also be advantageous to generate maps representing rates of statistical significance. Maps should serve as a clear, rapid and informative summary of complex geographical information and should help the reader identify interesting trends and generate relevant hypotheses.

10. Repeat the mapping analysis (step 9) using different normalized rates. Map the data in different formats and beware of "biased mapping" which was discussed elsewhere [44]. Ensure plotting maps that convey the message clearly and accurately.

11. Visualize and further analyze the plotted maps and note the presence of disease clusters ("disease hot-spots") as well as areas of significantly low incidence/mortality rates ("cold-spots"). Observe for interesting trends, particularly, if several of these clusters occur geographically side-by-side and are supported by hypotheses/current evidence of disease pathogenesis. It is often useful to compare generated disease maps with land-use maps that can be obtained from local authorities.

12. Perform sub-analysis of the identified "disease hot-spots" and correlate with the surrounding environment for any prevalent occupations, exposures, environmental factors, etc. If the patients within the area of high incidence (e.g. within a zip/postal code or a city) demonstrate an additional level of clustering (e.g., living on the same street or up and down the stream or river) it can further strengthen clustering findings and provide clues regarding possible triggers/exposures.

2.4.2. Limitations and bias

As illustrated in this chapter, studying the spatial patterns and geographical distribution of diseases has many benefits including the identification of disease clusters. This can be a powerful tool to help identify disease triggers and to better allocate financial and logistic resources for better management of these medical conditions. When the analysis is conducted properly, results are often specific. However, as in any type of analysis, one must be aware of potential limitations and intrinsic bias of the method. When analyzing clusters of patients in a given geographical region, one must be aware that there is a possibility that at least some of the observed clusters may be occurring by chance alone. Another important point, when studying the incidence of rare diseases in small regions: it is imperative to bracket the population analysis to at least 5000–10,000 residents per geographical area to reduce erroneous false-positive hits. Also, association does not always imply causality. Extensive additional field and experimental work must be performed to link identified associations causally with a given disease. Finally, one must be careful when directly comparing different geographic clustering studies as differences in the inclusion criteria, statistical methods or intrinsic differences of the populations at risk can produce divergent results.

3. Conclusions

The applications of cluster studies in medicine have developed rather rapidly in recent decades. These will enable us to focus on studying risk factors and possible etiologic triggers of rare cancers and other conditions. Furthermore, this work can help make informed decisions regarding resource allocation and promote the development of primary prevention programs.

Acknowledgements

The authors would like to sincerely thank both Dr. Linda Moreau and Dr. Elham Rahme for their generous support and valuable advice.

Conflict of interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this book chapter.

Author details

Feras M. Ghazawi[1], Steven J. Glassman[1], Denis Sasseville[2] and Ivan V. Litvinov[2]*

*Address all correspondence to: ivan.litvinov@mcgill.ca

1 Division of Dermatology, University of Ottawa, Ottawa, Ontario, Canada

2 Division of Dermatology, McGill University Health Centre, Montréal, Québec, Canada

References

[1] Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, et al. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. CA: A Cancer Journal for Clinicians. 2017;**68**(1):31-54. Epub 2017/11/22

[2] Elliott P, Wartenberg D. Spatial epidemiology: Current approaches and future challenges. Environmental Health Perspectives. 2004;**112**(9):998-1006 Epub 2004/06/17

[3] Rezaeian M, Dunn G, St Leger S, Appleby L. Geographical epidemiology, spatial analysis and geographical information systems: A multidisciplinary glossary. Journal of Epidemiology and Community Health. 2007;**61**(2):98-102 Epub 2007/01/20

[4] Torabi M, Rosychuk RJ. An examination of five spatial disease clustering methodologies for the identification of childhood cancer clusters in Alberta, Canada. Spatial and Spatio-Temporal Epidemiology. 2011;**2**(4):321-330 Epub 2012/07/04

[5] Halliday S. Death and miasma in Victorian London: An obstinate belief. BMJ. 2001;**323**(7327):1469-1471 Epub 2001/12/26

[6] Snow SJ. John snow: The making of a hero? Lancet. 2008;**372**(9632):22-23 Epub 2008/07/10

[7] Snow J. On the Mode of Communication of Cholera. 2nd ed. London,: J. Churchill. 1855;**vii**(1):162

[8] Wagner JC, Sleggs CA, Marchand P. Diffuse pleural mesothelioma and asbestos exposure in the north western Cape Province. British Journal of Industrial Medicine. 1960;**17**:260-271 Epub 1960/10/01

[9] McDonald JC. Epidemiology of malignant mesothelioma—An outline. The Annals of Occupational Hygiene. 2010;**54**(8):851-857 Epub 2010/11/10

[10] Geyer L. Über die chronischen Hautveränderungen beim Arsenicismus und Betrachtungen über die Massenerkrankungen in Reichenstein in Schlesien. Archiv für Dermatologie und Syphilis. 1898;**43**(1):221-280

[11] Schwartz RA. Reichenstein disease. International Journal of Dermatology. 1991;**30**(4):304-305 Epub 1991/04/01

[12] Bogle MA, Riddle CC, Triana EM, Jones D, Duvic M. Primary cutaneous B-cell lymphoma. Journal of the American Academy of Dermatology. 2005;**53**(3):479-484

[13] Suzuki R. Pathogenesis and treatment of extranodal natural killer/T-cell lymphoma. Seminars in Hematology. 2014;**51**(1):42-51

[14] Tsukasaki K, Tobinai K. Human T-cell lymphotropic virus type I-associated adult T-cell leukemia-lymphoma: New directions in clinical research. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research. 2014;**20**(20):5217-5225

[15] Gip L, Nilsson E. Ansamling av mycosis fungoides i Vasternorrlands lan [Clustering of mycosis fungoides in the county of Vasternorrland]. Lakartidningen. 1977;**74**(12):1174-1176 Epub 1977/03/23

[16] Litvinov IV, Tetzlaff MT, Rahme E, Jennings MA, Risser DR, Gangar P, et al. Demographic patterns of cutaneous T-cell lymphoma incidence in Texas based on two different cancer registries. Cancer Medicine. 2015;**4**(9):1440-1447 Epub 2015/07/03

[17] Litvinov IV, Tetzlaff MT, Rahme E, Habel Y, Risser DR, Gangar P, et al. Identification of geographic clustering and regions spared by cutaneous T-cell lymphoma in Texas using 2 distinct cancer registries. Cancer. 2015;**121**(12):1993-2003 Epub 2015/03/03

[18] Moreau JF, Buchanich JM, Geskin JZ, Akilov OE, Geskin LJ. Non-random geographic distribution of patients with cutaneous T-cell lymphoma in the greater Pittsburgh area. Dermatology Online Journal. 2014;**20**(7):pii: 13030/qt4nw7592w. Epub 2014/07/22

[19] Hazen PG, Michel B. Hodgkin's disease and mycosis fungoides in a married couple. Dermatologica. 1977;**154**(5):257-260 Epub 1977/01/01

[20] Hodak E, Klein T, Gabay B, Ben-Amitai D, Bergman R, Gdalevich M, et al. Familial mycosis fungoides: Report of 6 kindreds and a study of the HLA system. Journal of the American Academy of Dermatology. 2005;**52**(3 Pt 1):393-402 Epub 2005/03/12

[21] Litvinov IV, Shtreis A, Kobayashi K, Glassman S, Tsang M, Woetmann A, et al. Investigating potential exogenous tumor initiating and promoting factors for cutaneous T-cell lymphomas (CTCL), a rare skin malignancy. Oncoimmunology. 2016;**5**(7):e1175799 Epub 2016/09/14

[22] Ghazawi FM, Netchiporouk E, Rahme E, Tsang M, Moreau L, Glassman S, et al. Comprehensive analysis of cutaneous T-cell lymphoma (CTCL) incidence and mortality in Canada reveals changing trends and geographic clustering for this malignancy. Cancer. 2017;**123**(18):3550-3567 Epub 2017/05/12

[23] Ghazawi FM, Netchiporouk E, Rahme E, Tsang M, Moreau L, Glassman S, et al. Distribution and clustering of cutaneous T-cell lymphoma (CTCL) cases in Canada during 1992 to 2010. Journal of Cutaneous Medicine and Surgery. 2018 Mar/Apr;**22**(2): 154-165. DOI: 10.1177/1203475417745825. Epub 2017/12/16

[24] Zelen SWLBJWM. An analysis of contaminated well water and health effects in Woburn, Massachusetts. Journal of the American Statistical Society. 1986;**81**:583-596

[25] Burger M, Catto JW, Dalbagni G, Grossman HB, Herr H, Karakiewicz P, et al. Epidemiology and risk factors of urothelial bladder cancer. European Urology. 2013;**63**(2):234-241 Epub 2012/08/11

[26] Frumin E, Velez H, Bingham E, Gillen M, Brathwaite M, LaBarck R. Occupational bladder cancer in textile dyeing and printing workers: Six cases and their significance for screening programs. Journal of Occupational Medicine. 1990;**32**(9):887-890 Epub 1990/09/01

[27] Kurtzke JF. Multiple sclerosis in time and space—Geographic clues to cause. Journal of Neurovirology. 2000;**6**(Suppl 2):S134-S140 Epub 2000/06/29

[28] Bezzini D, Pepe P, Profili F, Meucci G, Ulivelli M, Bartalini S, et al. Multiple sclerosis spatial cluster in Tuscany. Neurological Sciences. 2017;**38**(12):2183-2187. Epub 2017/10/12

[29] Sheremata WA, Poskanzer DC, Withum DG, MacLeod CL, Whiteside ME. Unusual occurrence on a tropical island of multiple sclerosis. Lancet. 1985;**2**(8455):618 Epub 1985/09/14

[30] Schiffer RB, McDermott MP, Copley C. A multiple sclerosis cluster associated with a small, north-central Illinois community. Archives of Environmental Health. 2001;**56**(5):389-395 Epub 2002/01/05

[31] Jin Y, de Pedro-Cuesta J, Soderstrom M, Stawiarz L, Link H. Seasonal patterns in optic neuritis and multiple sclerosis: A meta-analysis. Journal of the Neurological Sciences. 2000;**181**(1-2):56-64 Epub 2000/12/02

[32] Simpson S Jr, Blizzard L, Otahal P, Van der Mei I, Taylor B. Latitude is significantly associated with the prevalence of multiple sclerosis: A meta-analysis. Journal of Neurology, Neurosurgery, and Psychiatry. 2011;**82**(10):1132-1141 Epub 2011/04/12

[33] Koch-Henriksen N, Sorensen PS. The changing demographic pattern of multiple sclerosis epidemiology. The Lancet Neurology. 2010;**9**(5):520-532 Epub 2010/04/20

[34] Russ TC, Batty GD, Hearnshaw GF, Fenton C, Starr JM. Geographical variation in dementia: Systematic review with meta-analysis. International Journal of Epidemiology. 2012;**41**(4):1012-1032 Epub 2012/07/17

[35] Vural H, Demirin H, Kara Y, Eren I, Delibas N. Alterations of plasma magnesium, copper, zinc, iron and selenium concentrations and some related erythrocyte antioxidant enzyme activities in patients with Alzheimer's disease. Journal of Trace Elements in Medicine and Biology. 2010;**24**(3):169-173 Epub 2010/06/24

[36] Chatin A. Recherches sur l'iode des eaux douces; de la presence de ce xorps sand les plantes at les animaux terrestes. Comptes Rendus de l'Académie des Sciences. 1852;**35**:505-517

[37] Leung AM, Braverman LE, Pearce EN. History of U.S. iodine fortification and supplementation. Nutrients. 2012;**4**(11):1740-1746 Epub 2012/12/04

[38] Pearce EN. National trends in iodine nutrition: Is everyone getting enough? Thyroid: Official Journal of the American Thyroid Association. 2007;**17**(9):823-827 Epub 2007/10/25

[39] Markel H. "When it rains it pours": Endemic goiter, iodized salt, and David Murray Cowie, MD. American Journal of Public Health. 1987;**77**(2):219-229 Epub 1987/02/01

[40] Markel H. A grain of salt. The Milbank Quarterly. 2014;**92**(3):407-412 Epub 2014/09/10

[41] Manjunath B, Suman G, Hemanth T, Shivaraj NS, Murthy NS. Prevalence and factors associated with goitre among 6-12-year-old children in a rural area of Karnataka in South India. Biological Trace Element Research. 2016;**169**(1):22-26 Epub 2015/06/13

[42] Gupta RK, Langer B, Raina SK, Kumari R, Jan R, Rani R. Goiter prevalence in school-going children: A cross-sectional study in two border districts of sub-Himalayan Jammu and Kashmir. Journal of Family Medicine and Primary Care. 2016;**5**(4):825-828 Epub 2017/03/30

[43] Wartenberg D. Using disease-cluster and small-area analyses to study environmental justice. In: Toward Environmental Justice: Research, Education, and Health Policy Needs. USA: National Academies Press; 1999:23-35

[44] Monmonier MS. How to Lie with Maps, Vol. xiii. 2nd ed. Chicago: University of Chicago Press; 1996. p. 207

Chapter 9

Real-Time Tele-Auscultation Consultation Services over the Internet: Effects of the Internet Quality of Service

Sinchai Kamolphiwong, Thossapon Kamolphiwong,
Soontorn Saechow and Verapol Chandeeying

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.74680

Abstract

A real-time tele-auscultation over the Internet is effective medical services that increase the accessibility of healthcare services to remote areas. However, the quality of auscultation's sounds transmitted over the Internet is the most critical issue, especially in real-time service. Packet loss and packet delay variations are the main factors. There is little knowledge of these factors affecting auscultation's sounds transmitted over the Internet. In this work, we investigate the effects of packet loss and packet delay variations, in particular, heart and lung sounds with auscultation's sound over the Internet in real-time services. We have found that both sounds are more sensitive to packet delay variations than packet loss. Lung sounds are more sensitive than heart sounds due to their timing interpretation. Some different levels of packet loss can be tolerated, e.g., 10% for heart sounds and 2% for lung sounds. Packet delay variation boundary of 50 msec is recommended. In addition, we have developed the real-time tele-auscultation prototype that tries to minimize the packet delay variation. We have found that real-time waveform of auscultation's visualization can help physician's confident level for sound interpreting. Some techniques for quality of service improvement are suggested, e.g., noise reduction and user interface (UI).

Keywords: tele-auscultation, e-stethoscope, e-health, tele-medicine, packet loss and delay variations, heart and lung sounds

1. Introduction

Quality of healthcare services in rural areas is a critical issue. Most developing countries are actively on improving the quality of healthcare services with the short and long term policy,

increasing healthcare staffs and implementing new technologies are the two wildly example policies [1–3].

Generally, people who live in rural areas receive healthcare services at primary care unit as the first choice. However, a rural healthcare unit has some limitations such as infrastructure, healthcare staff and good/advanced medical equipment. Thus, referral of patients to the secondary or tertiary care unit is used for solving these problems, which lead to cost on traveling and waste of time. The effective care at the primary care units is one of the significant keys, which can improve a quality of healthcare services in rural areas. Telemedicine is one of the main keys, and telemedicine applications can enhance the accessibility of healthcare services through the collaborations between primary care unit and cooperation unit [4–8].

Interactive consulting over the Internet is a cost-effective way between physician and specialist. The real-time applications such as Skype [9–12], Google Hangout [13, 14], FaceTime [15, 16], and WebRTC [17–19] can make good experience to participants like a face-to-face communication. However, the real-time consulting between physician and specialist requires some specific information and equipment which depend on consultation's topic such as stethoscope for listen the body sounds on auscultation process. However, when data are transmitted over the Internet, loss and delay can be occurred, especially its sensitivity in a real-time application. Packet loss [20] and packet delay variations [21] are the two main factors that reduce sound quality in real-time communication, e.g., body sounds from the stethoscope.

In our work, we investigate the effects of using e-stethoscope caused by packet loss and delay in the real-time tele-auscultation system over the Internet.

Significance: There is a little knowledge on the effects caused by packet loss and delay to the real-time tele-auscultation system over the Internet. This work deeply investigates these effects on lung and heart sounds, to see the impacts and factors that influence the quality of tele-auscultation services. We discuss what effects will happen, their causes, and results by varying a number of packet loss, delay variations, and types of lung and heart sounds. It is significantly different from human conversation sounds. We have suggested the percentage of packet loss and packet delay variations to meet some confident level of sound interpreting which can increase the success rate and effective outcomes.

The rest of chapter is organized as follows: Section 2 describes an overview of tele-auscultation including the differences between traditional auscultation and tele-auscultation, system compositions, and types of services. Section 3 presents our prototype system design and development. Section 4 shows the analysis result of the effects of packet loss and packet delay variations of heart and lung sounds over the Internet service. Finally, we conclude of our work in Section 5.

2. Overview of tele-auscultation

Tele-auscultation is a system for providing a remote auscultation to another place. The main challenge for this service is how to find the suitable mechanism to transmit the auscultation's sound over the Internet effectively with an acceptable quality of auscultation's sounds and to find other related supporting systems.

2.1. Differences between traditional auscultation and tele-auscultation

Auscultation is the medical method for sound listening inside the patient body, on detecting and identifying the abnormal sounds [22, 23]. A tele-auscultation provides the medical examination method similar to the traditional one, but the key of tele-auscultation that steps over the traditional auscultation is the mechanism for transmission the auscultation's sounds on long distance. Internet technology and electronic stethoscope are the keys for driven tele-auscultation service. Moreover, experiences on using tele-auscultation may differ from the traditional one, it is not face-to-face and may experience some delay of body sound sent from the remote site. This awareness should be raised when using this service. **Table 1** summarizes the differences between traditional stethoscope and e-stethoscope.

2.2. Compositions of tele-auscultation system

In our study, tele-auscultation can be summarized in three main compositions: stethoscope, client application, and application server. In terms of processes, they are: capturing, processing, transmission, and display. The system overview of tele-auscultation structure is shown in **Figure 1**.

In **Figure 1**, the first component is stethoscope. It is an instrument to capture the sounds in the body, and is used widely in auscultation process [24, 25]. Currently, an e-stethoscope (electronic stethoscope) is a new generation of auscultation's device. It can improve the quality of auscultation sounds [26], as summarized in **Table 1**, and it is a source for transmitting the sounds from the body to the next component. The next component, client application, is software that captures the sounds sent by e-stethoscope. There are two techniques of sound capturing: the modern one is receiving directly from e-stethoscope via the transmission module embedded in the device [27–30], and the older one is receiving from stethoscope's ear tip via audio input device such as microphone [31]. Moreover, the client application is the main processing function for the auscultation's sounds such as improving sound's quality, sounds volume control, encoding, filtering, waveform rendering, transmitting/receiving sound between cooperated components, and sound play out. The last component, application server, is a center for managing the

Properties	Electronic stethoscope	Acoustic stethoscope
Record and playback	Yes/no	No
Volume control & amplification	Yes	No
Noise reduction	Yes	No
Power supply	Batteries	No
Transmission technologies	Bluetooth, RF, USB	No
Chest piece	Single side tunable, button and screen	Single side tunable, dual side tunable (bell and diaphragm)
Tubing	Similar	
Headset		
Sound signal	Digital	Analog

Table 1. Comparison between electronic stethoscope and acoustic stethoscope.

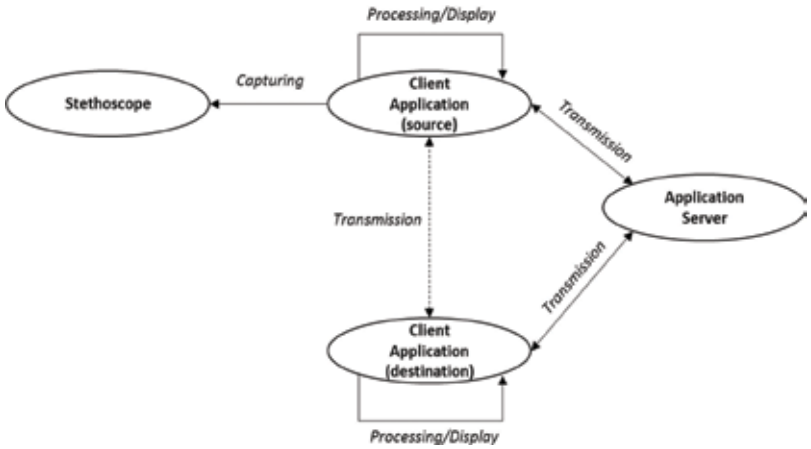

Figure 1. Overview of tele-auscultation architecture.

sessions, signaling, user accounts, forwarding the sounds (for client–server-service model). It should be noted that apart from client-server model, peer-to-peer service model may be considered. However, the server in this latter case may have less service requirements, e.g., for communication establishment, not for media streaming between client and server.

2.3. Types of tele-auscultation services

Synchronous and asynchronous (store-and-forward) are the two main communication types that describe the characteristic of each tele-auscultation service. Synchronous service is an interactive communication between participants [27–31], and auscultation' sound during the live session must be sent and played out immediately. While asynchronous service will store auscultation' sound in the middleware first, and then participants make a request and receive the data later [28, 29]. It seems that today technology with high speed Internet, asynchronous service may have small time delay while sending stored auscultation sounds to the remote site. However, since synchronous service is live session, physician can request different auscultation's sounds (from different body positions) from a patient or healthcare staff immediately to improve healthcare service level. For example, physician may ask healthcare staff to move the stethoscope up/down/left/right from the current position in place, to follow up the result from the sound interpretation, to capture a better sound quality (unclear sound). This will make the service quality of real-time tele-stethoscope much better over asynchronous one. However, the service requires some certain level of QoS, e.g., high speed link capacity. In conclusion, both services have different significant impacts to tele-auscultation services, and it depends on the purpose of usage and what practice scenarios are for of healthcare services.

2.4. Communication models

Client/server and peer-to-peer are the two widely used communication models that describe the characteristic of systems for sharing the media between source unit and destination unit.

In peer-to-peer model, it has no central server for managing the media stream. Each node directly communicates to each other. This will minimize the processing time at the central server as well as communication link time delay. Conversely, client/server model has a central server for managing almost everything, e.g., all information must be passed to the server first. However, for time-based information sharing, the client/server model will be useful. For tele-auscultation services, both models are used: client/server based model [29, 30], and peer-to-peer based model [28].

3. Design and development of a real-time tele-auscultation

We design and develop a real-time tele-auscultation application covering both communication models: peer-to-peer and client/server based models. The model consists of two main components. First component is client application that includes stethoscope controller, session controller, real-time audio waveform, and audio player. Another component is application server that comprises of account management and session management. The server is used for user authentication and session initiation between two client applications. Some more details are as follows:

3.1. Electronic stethoscope

In prototype demonstration, 3 M™ Littmann® Electronic Stethoscope Model 3200 [32] is utilized. The e-stethoscope provides a digital audio which in linear pulse-code modulation (LPCM) format, 4000 Hz of sampling rate and 16 bits per sampling.

As 4000 samples per second is not a standard voice sampling, we need to up-sample to 8000 times per second before putting in Real Time Transport Protocol (RTP) [34]. As a result, the sound quality is improved, but the bandwidth is increased to twice.

3.2. Auscultation's sound capturing

Stethoscope controller performs a connection to e-stethoscope by Bluetooth stack for capturing the auscultation's sound and handling some optional messages of e-stethoscope and other components. In our experiment, we captured the auscultation's sound every 50 msec, which created 400 bytes of audio data per packet. These audio packets will be transmitted to the destination node over the Internet for play-out at the destination.

3.3. Session and audio transmission

The handshaking mechanism is the important signal in peer-to-peer networks. Our prototype uses session initiation protocol (SIP) [33] and non-SIP (Web based signaling) for an initial protocol for real-time session establishing between participants. After completing the connection, the auscultation's sounds will be conveyed in RTP packets. There is no need for the central server participating in the media session.

4. Effects of packet loss and packet delay variations

In this section, study and analysis of auscultation's sounds quality affected by packet loss and packet delay variations in real-time communication over the Internet will be presented.

In our work, we focused and studied the characteristics of two body sounds: cardiac auscultation and lung auscultation. For the first one, cardiac auscultation, it is a method for screening heart sounds and heart murmurs [35]. In cardiac auscultation process, patient positions (left lateral recumbent, sitting and supine) and body locations (aortic area, pulmonic area, tricuspid area, and mitral area) can affect the quality of heart sounds. Particular positions and locations are important to the quality of listening to the specific heart sounds or heart murmurs [36–38]. Heart murmurs are the critical sounds related to the valvular heart disease. Intensity (grade), pitch, timing, and location have described the characteristic of each murmur [36–38]. For the second one, lung auscultation, it is a method for screening abnormal sounds over the lung's areas including front and back of the patient body [39]. Listening to the sounds of inspiration and expiration, together with comparing the intensity and pitch of each breath are the fundamental process for diagnosis the lung sounds [39, 40].

As noted above, a real-time auscultation service over the Internet should ensure a feasible and reliable. Packet loss and packet delay variations are the critical factors that significantly damage the heart and lung sound quality.

4.1. The packet loss and packet delay variation generator

We developed the software that can generate the difference level of packet loss and packet delay variation patterns, to study and analyze the effects to heart and lung sound quality. The software components are described in **Figure 2**.

- Sender: It sends auscultation's sound with a packet size of 400 bytes, 50 msec packet interval time. The following original three heart sounds [41] and five lung sounds [42] are observed, as shown in **Table 2**.

The waveforms of each sound are shown in **Figures 3** and **4**.

- Controller: It is a component for generating the patterns of packet loss and packet delay variations. The loss patterns were generated based on Gilbert-Elliot model [43–45] with 2, 5, 10, and 20% of loss values. The packet delay variation patterns were generated based on Poisson distribution [46] with 50, 60, and 70 msec time delays.
- Receiver: It is a component for converting the received packets to the play out format (LPCM, 4000 Hz, 16 bits, and 1 Mono-channel), with a controlled jitter buffer.

In this experiment, we measured the packet loss and delay variations between different network services. For example, Ethernet network was the Intranet in our experiment site which has a big bandwidth, the 3G/4G network was a service provided by local mobile operators, and ADSL in the remote area was the Internet connection provided by a local service provider. We collected test information for a week at different time. **Table 3** shows the average

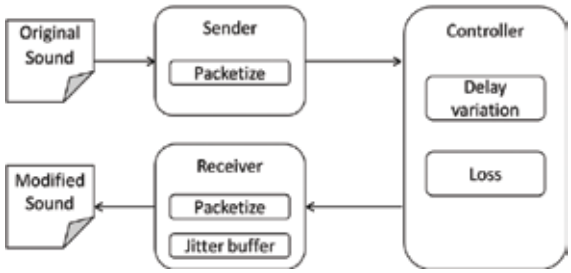

Figure 2. Software components for packet loss and packet delay variations generator.

Heart sounds	Lung sounds
1. Normal heart sound (75 bpm),	**1.** Coarse crackles (27 bpm),
2. Early systolic murmur (75 bpm),	**2.** Inspiratory stridor (23 bpm),
3. Pan-systolic murmur (75 bpm).	**3.** Normal vesicular (16 bpm),
	4. Pleural friction (19 bpm),
	5. Wheezing (27 bpm).

Table 2. The property of heart sounds and lung sounds.

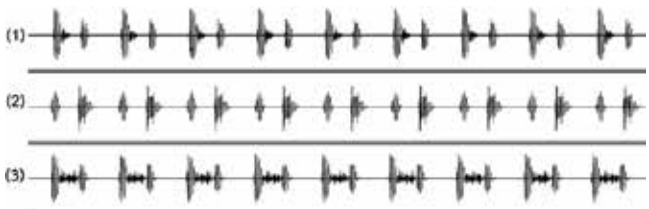

Figure 3. The original heart sounds (1) early systolic murmur, (2) heart normal, and (3) pan-systolic murmur.

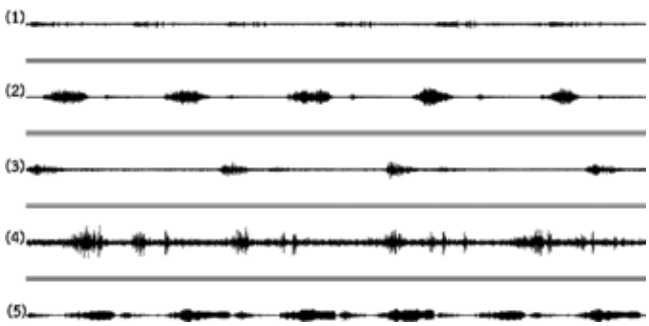

Figure 4. The original lung sounds (1) coarse crackles, (2) inspiratory stridor, (3) normal vesicular, (4) pleural friction, and (5) wheezing.

of packet loss and delays when information was sent across different network services. We noticed that the Intranet gave lowest packet loss and delays while ADSL in the remote area gave higher packet loss and larger delays.

Network types	Packet loss (%)		Packet delay variations (msec)	
	Range	Average	Range	Average
1) Sender: Ethernet network Receiver: Remote area via 3G/4G	0–3	2	10–70	55
2) Sender: Ethernet network Receiver: Ethernet network	0	0	0–5	2
3) Sender: Ethernet network Receiver: Remote area via ADSL	0–30	20	10–150	70

Table 3. Packet loss and delays between different network connection services.

4.2. Distortions of heart and lung sounds

Signal distortion is the alteration in the pulse of heart sound and the breath on lung sound. In our experiment method, we summarize the number of distortion pulses and breaths in a minute duration. Sample pulses of heart sound and breath of lung sound are shown in **Figures 5** and **6,** respectively.

4.2.1. Packet loss

In packet loss experiment, the heart and lung sound with 2, 5, 10, and 20% of packet loss are used.

Of replication five times, the signal distortions of three heart sounds and five lung sounds by comparing the pulse and breathing of each sound with its original sound. The following results are given:

Figure 7 shows the results of packet loss by varying from 2, 5, 10, and 20% for pan-systolic murmur. We can see that shape and position losses are randomly occurred from time to time where a higher value of packet loss gives more damage of shape and position. We tested all other given heart and lung sounds in **Table 2**.

Figure 8 shows sound damage positions of pan-systolic murmur randomly from five times of experiment where 20% of packet loss is applied. We can see that shape and position losses are randomly occurred from time to time. This will make the receiver node in hard condition for the result interpretation.

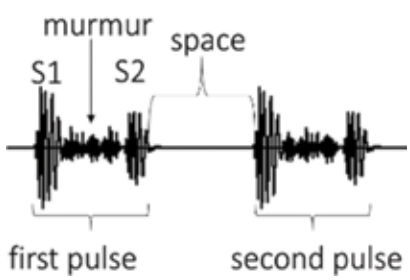

Figure 5. The pulse of hear sound.

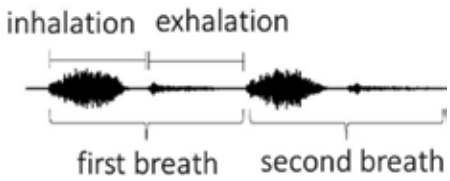

Figure 6. The breath of lung sound.

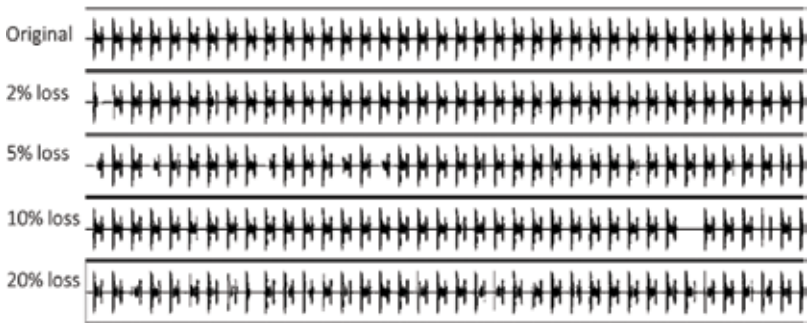

Figure 7. Packet loss varying from 2, 5, 10, 20% for pan-systolic murmur.

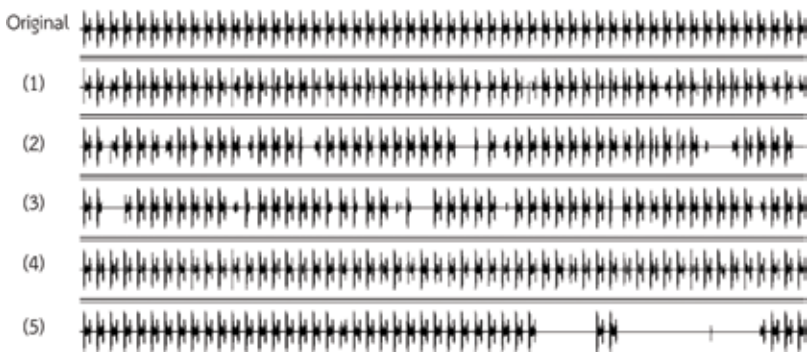

Figure 8. Randomly sound distortions and positions of pan-systolic murmur at 20% packet loss.

Of replication five times, with 2, 5, 10, and 20% of packet loss, the distortions are summarized in **Table 4**. The figures in the table are the percentage of distortions of each sound beats.

From **Table 4**, the increase of packet loss gets along with short-range distortions on heart sounds, but fluctuates and has long-range distortion of lung sounds. All of heart sounds and all of packet loss levels, the percent of distortions are less than 50%. On the other hand, for 2 and 5% of packet loss, the percent of distortions on lung sounds are less than 50%. However, when packet losses are 10 and 20%, the distortions go over 70%.

4.2.2. Packet delay variations

In packet delay variations experiment, the heart and lung sounds are investigated with three different levels of average packet delay variations; 50, 60, and 70 ms. Each delay was tested

Sounds	Packet loss			
	2%	5%	10%	20%
Heart sound				
Early systolic murmur	3	13	28	48
Heart normal	3	16	36	43
Pan-systolic murmur	6	24	27	48
Lung sound				
Coarse crackles	11	25	68	86
Inspiratory stridor	13	48	71	91
Normal vesicular	23	27	69	94
Pleural friction	25	46	68	98
Wheezing	20	37	79	85

Table 4. Percent of distortions among various heart and lung sounds on each level of packet loss.

for five times. **Figure 9** shows sample results of replication five times, the distortions of three heart sounds/five lung sounds by comparing the pulse and breathing of each sound with its original sound are shown in **Table 2**.

Of replication five times, the distortions of three heart sounds/five lung sounds by comparing the pulse and breathing of each sound with its original sound are shown in **Table 5**.

From **Table 5**, the packet delay variations at 50 msec get a short-range distortion on both heart and lung sounds and high distortion on both heart and lung sounds at 60 to 70 ms of packet delay variations. We can see that increasing a small delay variation, e.g., from 50 to 60 msec, significantly impact the sound distortion.

4.3. Evaluation of sound quality by assessor: packet loss

Ten medical professionals (they are medical doctors and nurses) who experience on auscultation operation at least 2 years participated in this evaluation. All of them are binding tests. Each person listens to three heart sounds/five lung sounds at packet loss of 2, 5, 10, and 20% without knowing which sound he or she is listening. According to the listening results, they

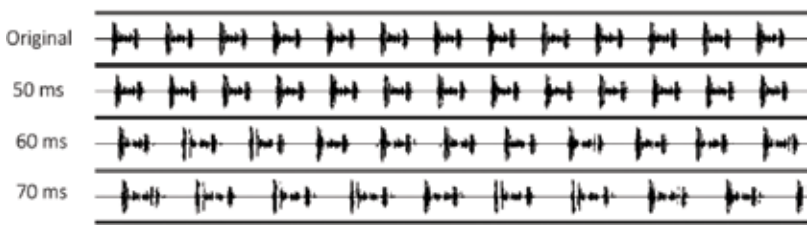

Figure 9. The result of packet delay variations for pan-systolic murmur.

Sounds	Packet delay variations (msec)		
	50	60	70
Heart sounds			
Early systolic murmur	3	83	100
Heart normal	3	72	100
Pan-systolic murmur	0	100	100
Lung sounds			
Coarse crackles	4	100	100
Inspiratory stridor	0	100	100
Normal vesicular	13	100	100
Pleural friction	5	100	100
Wheezing	15	100	100

Table 5. Percent of distortion among various heart and lung sounds on each level of packet delay variations.

indicated the type of sound with the confident level. The test result is shown in **Table 6**. It seems that most of them can detect normal heart, pan-systolic murmur, early systolic murmur, and normal vesicular sounds when a small packet loss is applied, e.g., less than 10%. More than 90% can correctly detect all heart sounds at 2–20% of packet loss. Percentage of correct detection on lung sounds depends on the type of lung sounds and level of packet loss.

4.4. Analysis of the effects of packet loss and packet delay variations

From the experiment of packet loss and packet delay variations, we analyzed the results as follows:

Sounds	Packet loss			
	2%	5%	10%	20%
Heart sounds				
Heart normal	100	100	100	90
Pan-systolic murmur	100	100	100	90
Early systolic murmur	100	90	90	90
Lung sounds				
Normal vesicular	100	100	90	60
Wheezing	100	90	80	30
Coarse crackles	70	60	40	70
Inspiratory stridor	90	80	50	70
Pleural friction	60	60	50	50

Table 6. Percent of correct detection on each sound.

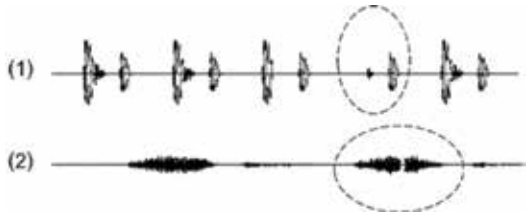

Figure 10. Sound discontinuity by packet loss, (1) early systolic murmur and (2) inspiratory stridor.

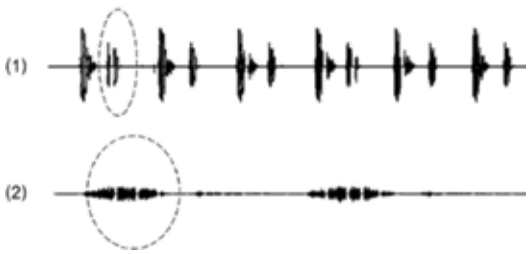

Figure 11. Sound discontinuity by packet delay variations, (1) early systolic murmur and (2) inspiratory stridor.

4.4.1. Sound missing and splitting

Auscultation method requires continued sound on each cycle to recognize the rhythm, pitch, and intensity. Packet loss and packet delay variations destroy the sounds continuity in different patterns randomly. Packet loss makes missing in some positions of sound, as shown in **Figure 10**, and packet delay variation makes split in some positions of sound, as shown in **Figure 11**.

Sound missing consists of two patterns: (1) either pulse or breath missing, and (2) some parts of pulse or breath missing, as shown in **Figure 12**. A pulse or breath missing is caused by burst loss on transmission process which is narrow range damage to heart and lung sounds. For another pattern, some parts of pulses or breaths missing, it is caused by uncertain pattern loss on transmission process. This pattern is wide range damage to heart and lung sounds. The increase of packet loss level cannot specify the pattern of sound discontinuity; it can only increase the missing sound.

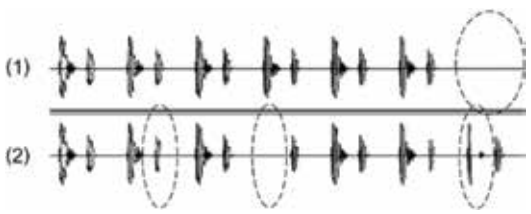

Figure 12. Behaviors of sound discontinuity on early systolic murmur after packet loss occurred, (1) a pulse and (2) some part of pulse.

http://dx.doi.org/10.5772/intechopen.74680

4.4.2. Pulse transformation

Transformation is the effect caused by packet loss. When the sound missing in some positions like murmur shape, it may lead to transformation to another type, and pulse transformation

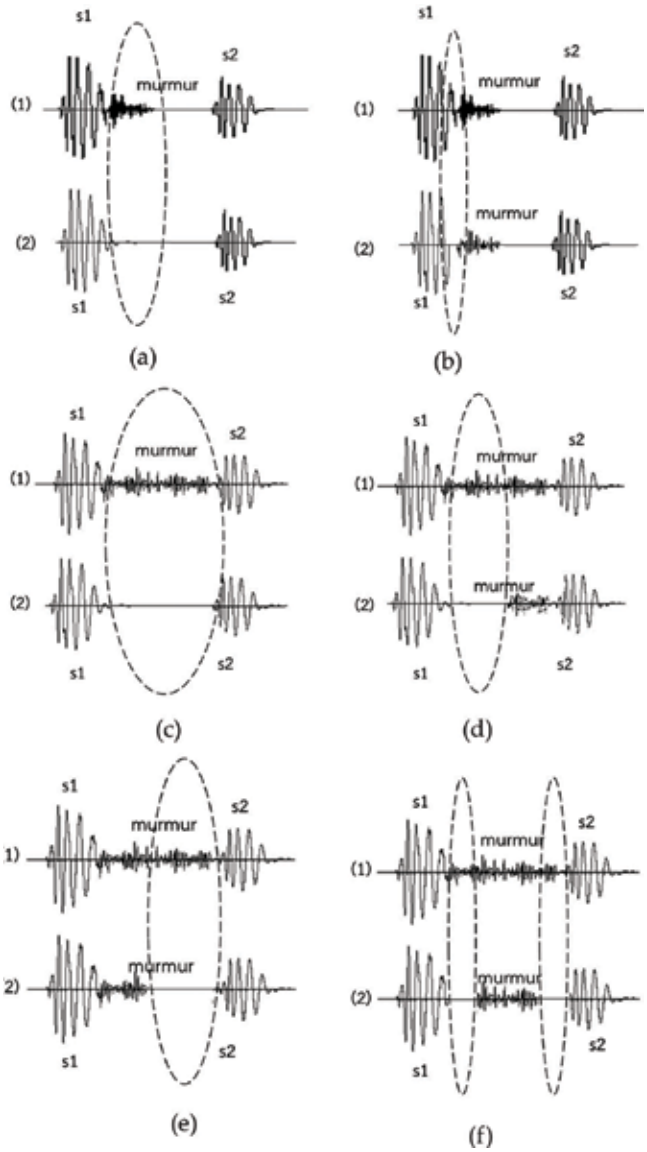

Figure 13. Samples of pulse transformations. (a) Early systolic murmur to heart normal (1) early systolic murmur (2) heart normal, (b) early systolic murmur to mid systolic murmur (1) early systolic murmur (2) mid-systolic murmur, (c) pan-systolic murmur to heart normal (1) pan-systolic murmur (2) heart normal, (d) pan-systolic murmur to late systolic murmur (1) pan-systolic murmur (2) late systolic murmur, (e) pan-systolic murmur to early systolic murmur (1) pan-systolic murmur (2) early systolic murmur, (f) pan-systolic murmur to mid systolic murmur (1) pan-systolic murmur (2) mid systolic murmur.

does not happen every pulse. The following result analysis of pulse transformation, as some examples, are as follows: Early systolic murmur sound transforms to normal heart sound, is a result of missing of the murmur shape but S1 and S2 are still remaining, as shown in **Figure 13(a)**. Transformation of early systolic murmur to mid-systolic murmur is a result of missing of the beginning part of murmur shape, as shown in **Figure 13(b)**. Transformation of pan-systolic murmur to normal heart sound is a result of missing a murmur shape, as shown in **Figure 13(c)**. Transformation of pan-systolic murmur to late systolic murmur is a result of missing of the first half of murmur shape, as shown in **Figure 13(d)**. Transformation of pan-systolic murmur to early systolic murmur is a result of missing of the second half of murmur shape, as shown in **Figure 13(e)**. Transformation of pan-systolic murmur to mid systolic murmur is a result of missing of the beginning part and tail of murmur shape, as shown in **Figure 13(f)**.

5. Design for improving the quality of service

Voice quality factors have been known for a long time, e.g., ITU guideline and standards [47], in tele-medicine applications, we do need to re-apply some techniques for this particular situation. The following design and implementation should be considered:

- ITU provides PLC (Packet loss concealment) technique for digital voice communications. However, waveform substitution (one technique in PLC) may not be appropriate. Zero insertion or silent insertion (another one in PLC) is more appropriate,
- Jitter buffer: jitter adaptation technique is deployed to reduce the effect of delay fluctuation due to the late or early arrival of voice packets (**Figure 14**). This will help the receiver-end hears sound in more comfortable level. However, due to real-time communication session condition, the buffer of delay absorption should be limited; we can have a small jitter buffer, e.g., few hundreds of milliseconds. In our experiment, 500 msec buffering (or 10 packets buffering) seems to be good enough. This will be traffic engineering choice to vary this figure.
- Noise removing: as mentioned, the device is operated remotely, and we noticed that moving of stethoscope creates a lot of noise. This creates an uncomfortable situation to the remote side. We have applied two stages noise filtering technique. The first stage looks at noise within a single voice packet (50 msec time interval), while the second stage evaluates the average noise energy for three consecutive packets. This will help a doctor in remote side working in convenience and comfortable way.

We have shown above that packet loss and delay will affect the quality of hearing, as a result, symptom determining may be hesitated. According to our prototype testing, most physicians at the remote side are happy with an e-stethoscope signal showing on the screen. This will help the interpretation confidence level after they are familiar with. We tested UI with packet loss and delay indicator, as shown in **Figure 15**, to raise awareness of a doctor during interacting operation. The levels of packet loss and delay can be easily noticed, for example, green color means there is no packet loss and delay (or very few, e.g., 1%), yellow color means there a few packet loss and delay (e.g., 5%), red color means there are high packet loss and delay

http://dx.doi.org/10.5772/intechopen.74680

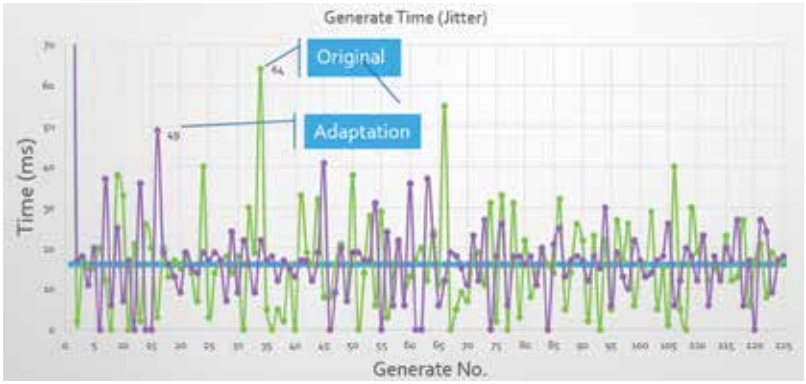

Figure 14. Jitter adaptation for time delay variation reduction.

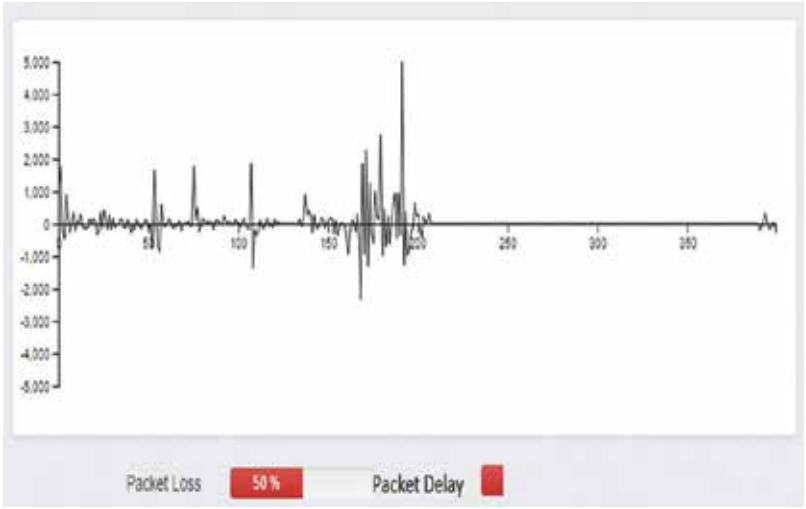

Figure 15. Real-time e-stethoscope signal with packet loss and packet delay indicators.

(e.g., more than 10%). We have concluded that with this UI designed, it helps for awareness of doctor's confidential level. Moreover, packet loss and delay pontificating levels can be adjusted according to the doctor experience.

6. Conclusion

This work studies the effects of packet loss and packet delay variations to real-time tele-auscultation services over the Internet. We categorize communications models of tele-auscultation services to asynchronous and synchronous, client/server and peer-to-peer. Some important compositions of tele-auscultation are drawn out with prototype software demonstration. We then focus and study the characteristics of two body sounds: cardiac auscultation and lung auscultation when these sounds are transmitted on the Internet in real-time applications.

From our experiment results, based on medical professional staff verification, we have found that both sounds are more sensitive to packet delay variations than packet loss. Lung sound is more sensitive than heart sound due to its timing interpretation, to recognize the rhythm, pitch, and intensity. Some different levels of packet loss can be tolerated for both sounds, e.g., 10% for heart sounds, 2% for lung sounds. However, packet delay variation boundary of 50 msec is recommended. Based on our analysis, sound missing and split, and pulse transformations are the two factors that affect the sound quality. The pulse transformation result may lead to misinterpreting of abnormal sounds. We have also found that making distinct normal sounds is more accurate than abnormal sounds. In addition to our prototype software, we have concluded that real-time waveform of auscultation's visualization can help physician confident level for sound interpreting. Moreover, showing the ratio of packet loss and delay variations in clear icon will raise awareness and increase the success rate and effective outcomes.

Acknowledgements

This work is supported by the Higher Education Research Promotion and National Research University Project of Thailand, Office of the Higher Education Commission (under the funding no. MED540548S at Prince of Songkla University).

Author details

Sinchai Kamolphiwong[1]*, Thossapon Kamolphiwong[1], Soontorn Saechow[1] and Verapol Chandeeying[2]

*Address all correspondence to: sinchai.k@psu.ac.th

1 Department of Computer Engineering, Faculty of Engineering, Prince of Songkla University, Hatyai, Songkla, Thailand

2 Faculty of Medicine, University of Phayao, Muang, Phayao, Thailand

References

[1] World Health Organization. Increasing Access to Health Workers in Remote and Rural Areas Through Improved Retention: Global policy recommendations. World Health Organization; 2010

[2] Dummer TJB, Cook IG. Exploring China's rural health crisis: Processes and policy implications. Health Policy. 2007;**83**(1):1-16

[3] Agyepong IA, Adjei S. Public social policy development and implementation: A case study of the Ghana National Health Insurance scheme. Health Policy and Planning. 2008;**23**(2):150-160

[4] World Health Organization. Telemedicine: Opportunities and Developments in Member States: Report on the Second Global Survey on eHealth. World Health Organization; 2010

[5] Kijsanayotin B, Kasitipradith N, Pannarunothai S. eHealth in Thailand: The current status. Studies in Health Technology and Informatics. 2010;**160**(Pt 1):376

[6] Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: A systematic review of reviews. International Journal of Medical Informatics. 2010;**79**(11):736-771

[7] Dávalos ME, French MT, Burdick AE, Simmons SC. Economic evaluation of telemedicine: Review of the literature and research guidelines for benefit–cost analysis. Telemedicine and e-Health. 2009;**15**(10):933-948

[8] Eron L. Telemedicine: The future of outpatient therapy?. Clinical Infectious Diseases. 2010;**51**(Supplement no 2):S224-S230

[9] Armfield NR, Bradford M, Bradford NK. The clinical use of Skype—For which patients, with which problems and in which settings? A snapshot review of the literature. International Journal of Medical Informatics. 2015;**84**(10):737-742

[10] Nield M, Hoo GW. Real-time telehealth for COPD self-management using skype™. COPD: Journal of Chronic Obstructive Pulmonary Disease. 2012;**9**(6):611-619

[11] Reynolds HN, Rogove H, Bander J, McCambridge M, Cowboy E, Niemeier M. A working lexicon for the tele-intensive care unit: We need to define tele-intensive care unit to grow and understand it. Telemedicine and e-Health. 2011;**17**(10):773-783

[12] Brecher DB. The use of Skype in a community hospital inpatient palliative medicine consultation service. Journal of Palliative Medicine. 2013;**16**(1):110-112

[13] Engle X, Aird J, Tho L, Bintcliffe F, Monsell F, Gollogly J, Noor S. Combining continuing education with expert consultation via telemedicine in Cambodia. Tropical Doctor. 2013:0049475513515654

[14] Lee JF, Schieltz KM, Suess AN, Wacker DP, Romani PW, Lindgren SD, Kopelman TG, Dalmau YCP. Guidelines for developing telehealth services and troubleshooting problems with telehealth technology when coaching parents to conduct functional analyses and functional communication training in their homes. Behavior Analysis in Practice. 2015;**8**(2):190-200

[15] Armstrong DG, Giovinco N, Mills JL, Rogers LC. FaceTime for physicians: Using real time mobile phone–based videoconferencing to augment diagnosis and care in telemedicine. Eplasty. 2011;**11**

[16] Brandt R, Hensley D. Teledermatology: The use of ubiquitous technology to redefine traditional medical instruction, collaboration, and consultation. The Journal of Clinical and Aesthetic Dermatology. 2012;**5**(11):35

[17] Cola C, Valean H. E-health appointment solution, a web based approach. In: Proceedings of E-Health and Bioengineering Conference (EHB), IEEE; 2015. p. 1-4

[18] Vidul AP, Hari S, Pranave KP, Vysakh KJ, Archana KR. Telemedicine for emergency care management using WebRTC. In: Proceedings of Advances in Computing,

Communications and Informatics (ICACCI), International Conference on, IEEE; 2015. pp. 1741-1745

[19] Jang-Jaccard J, Nepal S, Celler B, Yan B. WebRTC-based video conferencing service for telehealth. Computing. 2016;**98**(1-2):169-193

[20] Bolot Jean-Chrysotome. End-to-end packet delay and loss behavior in the internet. In: ACM SIGCOMM Computer Communication Review. Vol. 23, No. 4. New York, NY, USA: ACM; 1993. pp. 289-298

[21] Demichelis C, Chimento P. "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM); 2002

[22] Dorland's Medical Dictionary for Health Consumers. S.v. "auscultation". Available from: http://medical-dictionary.thefreedictionary.com/auscultation [Accessed: February 10, 2016]

[23] The American Heritage® Medical Dictionary. S.v. "auscultation". Available from: http://medical-dictionary.thefreedictionary.com/auscultation, [Accessed: February 10, 2016]

[24] Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health. 7th edition. S.v. "stethoscope". Available from: http://medical-dictionary.thefreedictionary.com/stethoscope. [Accessed: February 10, 2016]

[25] McGraw-Hill Concise Dictionary of Modern Medicine. S.v. "stethoscope". Retrieved February 10 2016 from: http://medical-dictionary.thefreedictionary.com/stethoscope

[26] Myint WW, Dillard B. An electronic stethoscope with diagnosis capability. In System Theory, 2001. In: Proceedings of the 33rd Southeastern Symposium on, IEEE; 2001. pp. 133-137

[27] Johanson M, Gustafsson M, Johansson L-A. A remote auscultation tool for advanced home health-care. Journal of Telemedicine and Telecare. 2002;**8**(2):45-46

[28] Xu, Lisheng, Ying Wang, Yue Wang, Ning Geng, Yao Jiang, Guanxiong Wang, Jiajin Liu, and Cong Feng. "The design and implementation of telemedical consulting system for auscultation." In: Proceedings of 2011 IEEE International Conference on Information and Automation, IEEE; 2011. pp. 242-247

[29] McMechan, Christian, Irina Morozov, Aaron Patten, and Poman So. "Tele-auscultation system." In: Proceedings of Broadband and Wireless Computing, Communication and Applications (BWCCA), 2011 International Conference on, IEEE; 2011. pp. 478-481

[30] Foche-Perez I, Ramirez-Payba R, Hirigoyen-Emparanza G, Balducci-Gonzalez F, Simo-Reigadas F-J, Seoane-Pascual J, Corral-Peñafiel J, Martinez-Fernandez A. An open real-time tele-stethoscopy system. BioMedical Engineering. 1186;**11**(2012):57

[31] Lu B-Y, Hsu L-Y, Wu H-D, Hsueh M-L, Sing S-S, Tang R-H, Su M-J, Wang J-C, Lai J-S. Real-time mobile-to-mobile stethoscope for distant healthcare. In: Proceedings of Advanced Communication Technology (ICACT), 2014 16th International Conference on, IEEE. 2014. pp. 151-156

[32] 3M™ Littmann® stethoscope. Electronic Stethoscope Model 3200. Internet: http://solutions.3m.com/wps/portal/3M/en_US/Littmann/stethoscope/electronic-auscultation/model-3000-series/ [Accessed: June 10, 2015]

[33] Rosenberg, Jonathan, Henning Schulzrinne, Gonzalo Camarillo, Alan Johnston, Jon Peterson, Robert Sparks, Mark Handley, and Eve Schooler. SIP: Session Initiation Protocol. Vol. 23. RFC 3261, Internet Engineering Task Force; 2002

[34] Jacobson V, Frederick R, Casner S, Schulzrinne H. RTP: A transport protocol for real-time applications. IETF RFC. 2003:3550

[35] Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health. 7th Edition. S.v. "cardiac auscultation". Available from http://medical-dictionary.thefreedictionary.com/Cardiac+auscultation [Accessed: June 10, 2015]

[36] Tavel ME. Cardiac auscultation a glorious past—And it does have a future! Circulation. 2006;**113**(9):1255-1259

[37] Chizner MA. Cardiac auscultation: Rediscovering the lost art. Current Problems in Cardiology. 2008;**33**(7):326-408

[38] Mangione S, Nieman LZ, Gracely E, Kaye D. The teaching and practice of cardiac auscultation during internal medicine and cardiology training: A nationwide survey. Annals of Internal Medicine. 1993;**119**(1):47-54

[39] Bohadana A, Izbicki G, Kraman SS. Fundamentals of lung auscultation. New England Journal of Medicine. 2014;**370**(8):744-751

[40] Scott PR. Lung auscultation recordings from normal sheep and from sheep with well-defined respiratory tract pathology. Small Ruminant Research. 2010;**92**(1):104-107

[41] 3M™ Littmann® stethoscope. Basic Heart Sounds Course. Available from: http://www.littmann.ca/wps/portal/3M/en_CA/3M-Littmann-CA/stethoscope/littmann-learning-institute/heart-lung-sounds/heart-sounds/. [Accessed: June 10, 2015]

[42] 3M™ Littmann® stethoscope. Listen to Lung Sounds. Available from: http://www.littmann.ca/wps/portal/3M/en_CA/3M-Littmann-CA/stethoscope/littmann-learning-institute/heart-lung-sounds/lung-sounds/. [Accessed: June 10, 06]

[43] Gilbert EN. Capacity of a burst-noise channel. Bell System Technical Journal. 1960;**39**(5):1253-1265

[44] Elliott EO. Estimates of error rates for codes on burst-noise channels. Bell System Technical Journal. 1963;**42**(5):1977-1997

[45] Haßlinger G, Hohlfeld O. The Gilbert-Elliott model for packet loss in real time services on the internet. In: Proceedings of Measuring, Modelling and Evaluation of Computer and Communication Systems (MMB), 14th GI/ITG Conference. 2008. pp. 1-15

[46] Saaty TL. Elements of Queueing Theory: With Applications. New York: McGraw-Hill; 1961

[47] ITU-T Recommendations G.114 (05/2003). Available from: https://www.itu.int/rec/T-REC-G.114/en [Accessed: June 10, 2015]

Chapter 10

Exploring the Interrelationship of Risk Factors for Supporting eHealth Knowledge-Based System

Geletaw Sahle Tegenaw

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.75033

Abstract

In developing countries like Africa, the physician-to-population ratio is below the World Health Organization (WHO) minimum recommendation. Because of the limited resource setting, the healthcare services did not get the equity of access to the use of health services, the sustainable health financing, and the quality of healthcare service provision. Efficient and effective teaching, alerting, and recommendation system are required to support the activities of the healthcare service. To alleviate those issues, creating a competitive eHealth knowledge-based system (KBS) will bring unlimited benefit. In this study, Apriori techniques are applied to malaria dataset to explore the degree of the association of risk factors. And then, integrate the output of data mining (i.e., the interrelationship of risk factors) with knowledge-based reasoning. Nearest neighbor retrieval algorithms (for retrieval) and voting method (to reuse tasks) are used to design and deliver personalized knowledge-based system.

Keywords: knowledge-based system, eHealth, pattern discovery, data mining, association rule

1. Introduction

In Africa, on average there are nine hospital beds per 10,000 people in comparison to the world average of 27. In sub-Saharan Africa, the physician-to-population ratio is the lowest in the world [1, 2]. Countries like Ethiopia set a strategic plan to improve access and equity to preventive, essential health interventions at the village and household levels to ensure healthcare coverage in rural areas [3, 4].

On the one hand, because of the limited resource setting, the healthcare services have not got the equity of access to the use of health services, the sustainable health financing, and the

quality of healthcare service provision. The physician-to-population ratio is below the World Health Organization (WHO) minimum recommendations [1, 2]. Still, pneumonia, diarrhea, acute upper respiratory tract infection, acute febrile illness, and malaria account for 64% under five morbidity [5].

> *About 68% of the country's total population living in areas at risk of malaria, 75% of the country is vulnerable to malaria (defined as areas <2000 m, those areas are fertile and suitable for agriculture and accounts for up to 17% of outpatient consultations, 15% of admissions and 29% of inpatient deaths.* [6, 7]

On the other hand, it has been more than four decades the computer program reasons (e.g., MYCIN was developed in the early 1970s) and uses knowledge to assist the domain experts and minimize routine activities. To prevent and control the crisis of malaria, different scholars and responsible bodies have made remarkable efforts by conducting researches, implementing strategies, and policies [6, 7]. A predictive data mining model has been constructed using Ethiopian WHO malaria database, metrological database, and national mapping database [8]. The model is accurate to determine the occurrence of death, and it is good enough to identify the cases.

However, the system is looking a mechanism to assist routine healthcare activities, administrative and medical cost, demographic challenges, and equitable health distribution. For instance, the health extension workers (HEWs) assist peripheral health services by bridging the gap between the communities and health facilities [3, 4]. Each kebele has two HEWs responsible for providing outreach services. A kebele is the smallest governmental administrative unit and on average has a population of 5000 people. The HEWs teach the community house to house for each and every person in the kebele in order to create and promote healthy lifestyles. In all, the healthcare system is searching a mechanism for teaching, alerting, and recommendation system to support the daily routine activities.

To alleviate those issues, creating a competitive eHealth knowledge-based system (KBS) is the main goal of this work which will bring unlimited benefit in low resource setting. As a case study, we choose malaria (malaria dataset) because malaria prevention and control at the community level face numerous challenges because of the climate condition (temperature, rainfall), epidemiological, and genetic, poverty, malaria outbreak, over prescription for positive result and so on. Knowing the pattern and interrelationship of risk factors is important for supporting knowledge-based system as well as prediction of malaria death occurrences/cases. An attempt is made for exploring the degree of association between malaria risk factors (related to the malaria death occurrence and case identification). Investigating the degree of interrelationship among risk factors will have a great contribution toward eradicating the outbreak of malaria. The outcome of the study helps to mitigate the severity through investigating the association of risk factors and building of a competitive knowledge-based system.

2. Literature review

Knowledge-based systems is aimed to understand or bring human-level intelligence through simulating or acting one or more of intelligent behaviors (such as thinking, problem-solving,

learning, understanding, emotions, consciousness, intuition and creativity, language capacity, etc.). On the one hand, KBS is advantageous when there is shortage of expert, decision-making for problem-solving needs an intelligent assistant, expertise is needed to be stored for future use, and so on. On the other hand, KBS faces a lot of challenges due to the abstract nature of knowledge, limitation of cognitive science, and other scientific methods [9, 10].

Knowledge representation and inference engine are the two building blocks of KBS. The knowledge acquired from experts, documents, books, and other resources has been organized using knowledge representation. The inference engine gets the knowledge and instructs how to use the knowledge to solve problems using rule- or case-based reasoning. Rule-based reasoning is a technique that reasons out about a problem based on the knowledge that is represented in the form of rules [11]. Case-based system represents situations or domain knowledge in the form of cases, and it uses case-based reasoning technique to solve new problems or to handle new situations [12].

Knowledge-based system (in health and medical domain) has made a remarkable effect through providing a reliable diagnostic and cost-effective service. Several systems have been implemented in different medical areas like cancer therapy, infections, blood diseases, general internal medicine, glaucoma, and pulmonary function tests [13, 14]. Such systems can be designed to exhaustively consider all possible diseases in a domain, which could outperform human experts to achieve a rapid and accurate diagnosis. Integrating and updating domain knowledge with knowledge discovery are relevant to increase the interestingness and user belief (such as matching discovered pattern with existing knowledge) [15]. Integrated eHealth knowledge-based system based on acquiring health knowledge will support users in exchange of knowledge and accessibility for the users through data collection, care documentation, and knowledge extraction [16].

Following and implementing a hybrid (integrated) intelligent system for medical data classification is good to produce effective knowledge-based system [17]. A promising result is scored through integration to improve the quality of knowledge-based system [17, 18]. For instance, integrating the result (rule) of the PART classification algorithm with the knowledge-based system has delivered a favorable result for the diagnosis and treatment of visceral leishmaniasis [19]. The paper by Seera and Lim has experimented and used fuzzy min-max neural networks to learn incrementally from sample data, classification, and regression tree for prediction and random forest model to achieve high classification performance [17]. Kerdprasop and Kerdprasop also tried to automate data mining model by focusing on post data mining process to step of automatic knowledge deployment using induced knowledge and formalization classification rules [20].

However, the need to work more in providing explanatory rules and handling missing data in real-world application is expected. A mechanism for handling irrelevant rule (result) is required in case of inductive experiment system and so on. To alleviate those issues, in our case study, we tried to explore the pattern and the interrelationship of risk factors for supporting the knowledge-based system as well as prediction of malaria death occurrences/cases. The result will increase the interestingness and belief of eHealth knowledge-based system which will bring unlimited benefit in low resource setting.

2.1. Research aim and objectives

Investigating the potential applicability of exploring the interrelationship of risk factors using data mining to create a competitive eHealth knowledge-based system is the main goal of this work.

2.2. Methodology

Cross industry standard process for data mining (CRISP-DM) methodology is adopted to investigate the interrelationship of malaria risk factors. Then, to design eHealth knowledge-based system, nearest neighbor retrieval algorithms (for retrieval) and voting method (to reuse tasks) are used. The technique is easy in exploring relevant cases and provides an opportunity to retrieve partially matching cases [21–24].

The malaria data is collected from a zonal health facility in each of the 86 zones of Ethiopia. To understand the problem domain, we used observation, interviewing with experts and data managers and reviewing documents, reports, and literatures. This helps us to select and integrate decisive attributes from different sources. The data selected from the WHO (World Health Organization) database is integrated with the decisive attributes (like temperature, rainfall, and altitude) extracted from the Ethiopian National Meteorological Agency and Mapping Agency in order to find the association of risk factors.

Exploratory data analysis is performed to get familiar with the data and prepared for investigating the degree of interrelationship. The data mining task is finding the internal association between data elements that will determine the occurrence of death/case. To maintain the quality of the data, preprocessing tasks such as data cleaning (handling missing values, noisy, and outer values), data integration tasks, and data transformation tasks are performed.

The collected Ethiopian WHO malaria database contains five basic attributes (more than 37, 000 records) that provide information about the geographic location and period of coverage. These attributes include country name, region (administrative regions from which malaria information is collected), zone and health facility name, year, and month. The detailed information of the attributes is categorized based on the WHO standards and explicitly represents the detailed information about malaria in each zone of the region across Ethiopia. These categories contain age (less than, equal, or greater than 5 year), malaria type (*P. vivax* and *P. falciparum*), cases (inpatient and outpatient), inpatient cases (cases and deaths), severe anemia (inpatient malaria cases less than 5 years and greater than 5 years), and uncomplicated lab-confirmed malaria less than 5 years and greater than 5 years (*P. vivax* outpatient cases and *P. falciparum* outpatient cases). Each attribute is preprocessed and statistically summarized into address, patient profile, weather, and altitude. For example, **Table 1** presented the statistical summary of uncomplicated malaria less than 5 years of lab-confirmed *Plasmodium falciparum*.

In order to extract hidden patterns and relationships within the data, from the initial dataset, a number of attributes are constructed. As shown in **Table 2**, from malaria with severe anemia, attributes such as age, malaria type, cases, and malaria visits, as well as the number of cases and deaths, are constructed.

Summary of the datasets compiled for association rule discovery with their possible nominal values and description is depicted in **Figure 1**. The malaria dataset used for this study consists

Uncomplicated Malaria <5 years P.FALCIPARUM Lab Confirmed				
Cases	Yes	No	Missing Value	Noisy Value
Frequency	105	4555	43	1
Percent	2.232%	96.832%	0.914%	0.021%

Table 1. Summary of uncomplicated malaria less than 5 years.

Initial Dataset Attribute	Target Dataset (Attribute Constructed)
Inpatient malaria with severe anemia < 5 years cases	- Age - Type of cases - Type of malaria - Type of malaria visits - Number of cases - Number of deaths
Inpatient malaria with severe anemia < 5 years deaths	
Inpatient malaria with severe anemia > 5 years cases	
Inpatient malaria with severe anemia >5 years deaths	

Table 2. Malaria with severe anemia and list of attributes constructed.

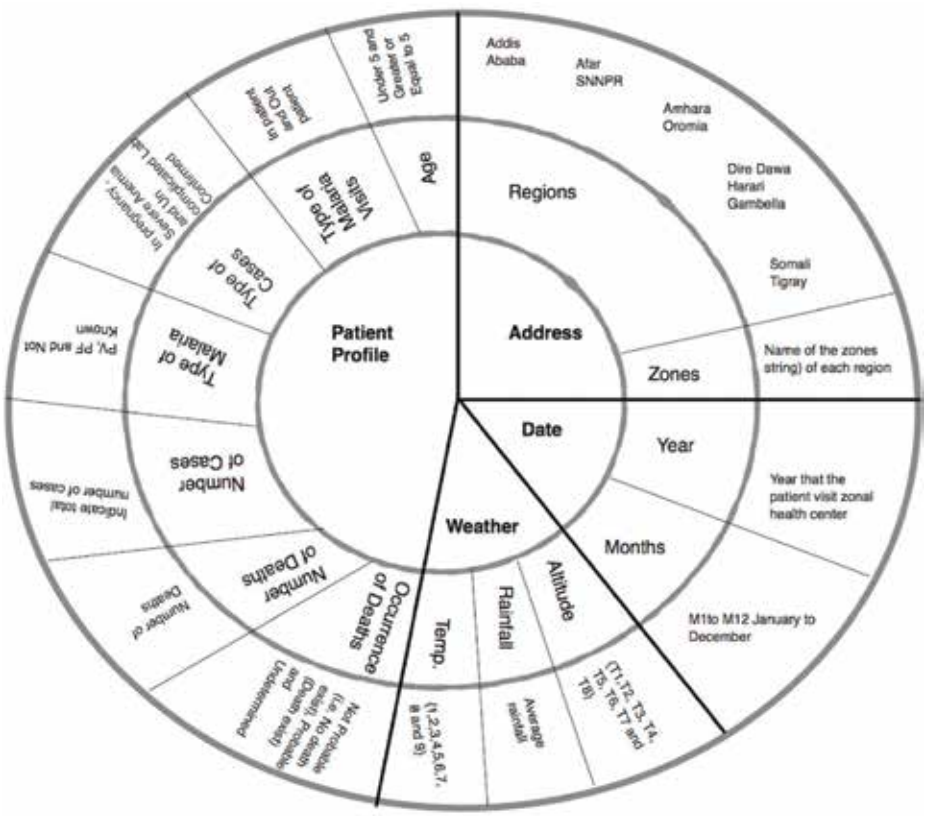

Figure 1. Summary of attributes and their transformed values.

of 14 attributes. The first inner part mainly presents the category of the attributes (i.e., profile, weather, address, and date), the second inner part presents the list of attributes constructed from category, and the last inner part indicates the value of each attribute. The region and zones contain all administrative regions in Ethiopia and the locations where the patients live. The date (year and month) indicates the year and the months that the patients visit the zonal health center. The age indicates the category of patients usually classified under 5 and greater or equal to 5. The type of malaria visits is classified into inpatient and outpatient malaria visits. The attribute type of case indicates the category of cases called in pregnancy, severe anemia, and uncomplicated lab-confirmed malaria. The type of malaria attributes contains PV, PF, and not-known values. In case of severe anemia and in pregnancy, the type of malaria is not known or determined in the dataset. A number of cases and deaths indicate the total number of cases and deaths, respectively. The attribute occurrence of deaths contains not probable (i.e., no deaths exist), probable (deaths exist), and undetermined values. The rainfall attribute contains a numeric value to represent average rainfall. In the case of outpatient visit, in pregnancy and uncomplicated lab-confirmed malaria, death is not known or listed in the dataset. The transformed temperature values 1, 2, 3, 4, 5, 6, 7, 8, and 9 represents 0–5^{0c}, 5–10^{0c}, 10–15^{0c},16–20^{0c}, 21–25^{0c}, 25–30^{0c}, 31–35^{0c}, 35–40^{0c}, and >40^{0c}, respectively. The transformed altitude values T1,T2, T3, T4, T5, T6, T7, and T8 represent >3500 m, 2500–3500 m, 2000–2500 m, 1500–2000 m, 1000–1500 m, 500–1000 m, 0–500 m, and <zero values, respectively.

3. Apriori method

The Apriori algorithm (a well-known association rule discovery method) takes a dataset with a list of items that can be easily transformed into a transaction form by creating an item for each attribute value pair that exists in the dataset [25–27]. Minimum support and minimum confidence thresholds are also defined to enable Apriori algorithms identify frequent items that are strongly associated. **Table 3** presented the step-by-step procedure to mine and extract frequent items using Apriori methods.

Given a support threshold (S), sets of X items that appear in greater than or equal to S baskets are called frequent item sets. Find all rules on item sets of the form X→Y with minimum support and confidence. For example, if-then rules about the content of the baskets {i1, i2,…,ik} → j means "if a basket contains all of i1,…,ik then it is likely to contain j." A typical question of Apriori is to "find all association rules with support ≥ S and confidence ≥ C." In

- *Initially, scan database once to get frequent 1-itemset*
- *Generate candidate k+1 itemsets from frequent k itemset. For each k, we construct two sets of k –tuples:*
  - C_k *= the set of candidate k – tuples*
  - L_k *= the set of truly frequent k –tuples.*
- *Test the candidates against DB*
- *Terminate when no frequent or candidate set can be generated*

Table 3. Apriori methods.

general, support of an association rule is the frequency of occurrence of the set of items it mentions, and confidence of this association rule is the probability of j given i1,…,ik. It is the number of transactions with i1,…,ik containing item j. This will measure the strength of associations between i1, i2,…,ik, and j.

The key concepts are frequent item sets (the set of items which have minimum support, denoted by Li for ith-item set), a priori property (any subset of frequent item set must be frequent), and join operation (to find Lk, a set of candidate k-item sets are generated by joining Lk-1 with itself). Once frequent item sets are obtained, it is straightforward to generate association rules with confidence larger than or equal to a user-specified minimum support and minimum confidence. The next top quality of the Apriori algorithm is to implement its achievement of good performance by reducing the size of candidate sets that are considered and selected for frequent k-item set [28].

A class implementing an Apriori-type algorithm iteratively reduces the minimum support until it finds the required number of rules with the given minimum confidence [29].

For mining Weka (Waikato Environment for Knowledge Analysis), knowledge discovery tool using Java is used. In Weka 3.7.3, if class association rule (car) property is enabled, the class association is mined instead of (general) association rules. Class classification generates rules that are frequently happening to the probable occurrence of malaria cases. In many studies, associative classification has been found to be more accurate than some traditional classification methods, such as C4.5 [25]. Associative classification can search strong associations between frequent patterns (conjunctions of attribute-value pairs) and class labels. Because association rules explore highly confident associations among multiple attributes, this approach may overcome some constraints introduced by decision tree induction, which considers only one attribute at a time. In all, in association of rule mining, finding all the rules that satisfy both a minimum support and a minimum confidence threshold is important so as to generate strong and interesting rules from the frequent patterns.

4. eHealth knowledge-based system

Knowledge-based systems are computer programs that try to solve problems in a human expert-like fashion by using knowledge about the application domain (knowledge base) and problem-solving techniques (inference method). The rule-based reasoning technique can be used with other reasoning techniques in order to make a knowledge-based system more efficient. For example, case-based and rule-based reasoning can be used together. Rule-based system is an example of knowledge-based system that uses rules for knowledge representation and rule-based reasoning for reasoning techniques. The development of knowledge-based systems in medical areas has made it possible to provide reliable and thorough diagnostic services with a minimum cost. Such systems can be designed to exhaustively consider all possible diseases in a domain, which could outperform human experts to achieve a rapid and accurate diagnosis. Several systems have been implemented in different medical areas like cancer therapy, infections, blood diseases, general internal medicine, glaucoma, and pulmonary function tests [13, 14].

Figure 2 presented the detail architecture of the proposed system. In this research we tried to integrate the output of data mining (i.e., the interrelationship of risk factors) with a knowledge-based system. Apriori algorithm using CRISP-DM methodology is adopted to create the interrelationship of risk factors and used for knowledge acquisition to develop a knowledge-based system. Nearest neighbor retrieval algorithms (for retrieval) and voting method (to reuse tasks) are used to design the eHealth knowledge-based system. The knowledge is represented using "IF a certain situation holds, THEN take a particular action," and the knowledge acquired (interrelationship of risk factors) from Apriori algorithm are rules. An

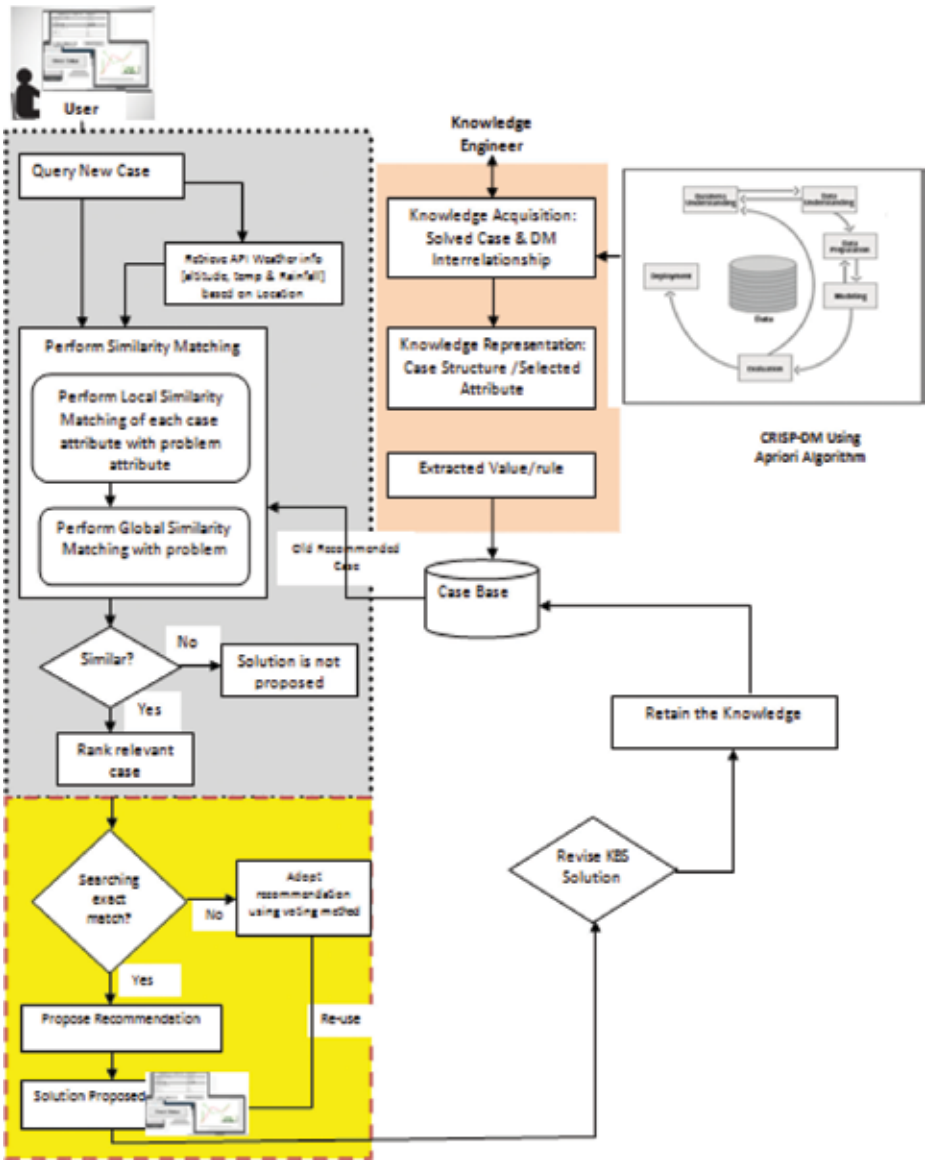

Figure 2. Adopted and proposed architecture for supporting eHealth KBS [19, 23, 24].

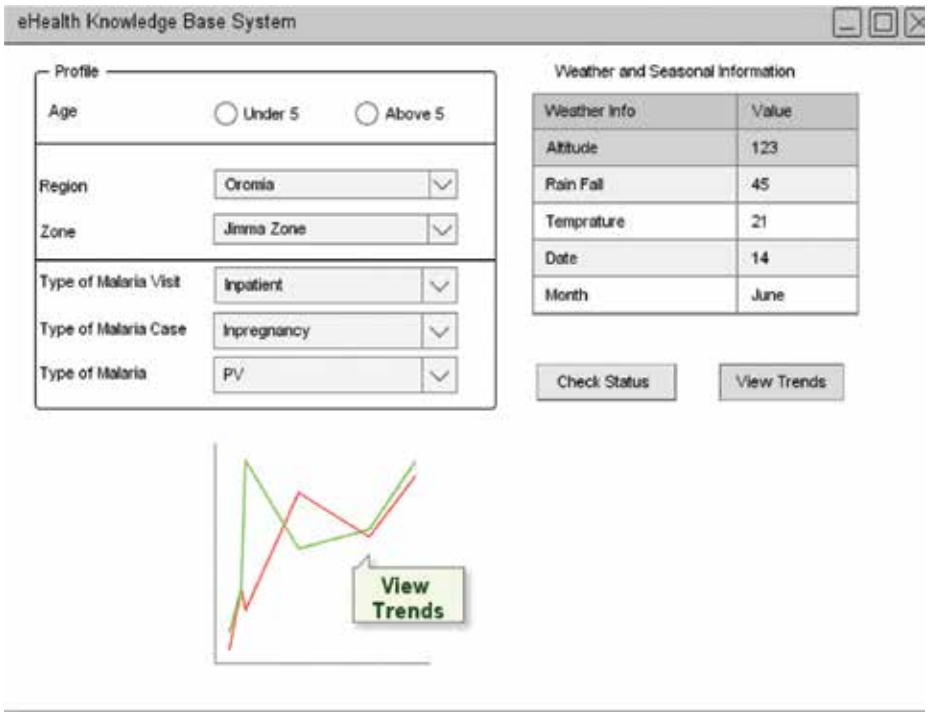

Figure 3. Graphical user interface.

inference engine of the knowledge-based system can derive the conclusion by looking the possible scenario and recommendations. The goal of the case study is to provide a personalized knowledge-based system solution using the support of interrelationship of risk factors. The user interfaces provide a communication between the user and the system.

Figure 3 presents the graphical user interface. The user tries to use the system or initiate queries by selecting profile and address information. Based on the desired location (using region and zone), weather information such as altitude, rainfall, and temperature are filled automatically from external weather API. Then, similarity matching is performed using the new queries to retrieve and recommend the proposed solution. However, if similarity matching is unsuccessful, voting technique is applied to select the relevant cases. Finally, to select or recommend the solution, the domain expert will evaluate and validate the new case. The knowledge-based system will use the validated case for future purpose.

5. Experimental results and discussions

The study tried to explore the interrelationship of risk factors for supporting eHealth knowledge-based system. We have used malaria dataset as a case study to discover the association among the various malaria risk factors using associative rule discovery data mining, and we integrate it with the eHealth knowledge-based system.

5.1. Experimental setup

A general and a class association rules are used to discover interesting association patterns. A total of 120 experiments executed using Apriori algorithm (60 experiments using general association rule and 60 experiments for class association rule mining) as depicted in **Table 4**. The confidence level is the most important parameter to attain the required objective. By considering this, the experiment is done at different confidence levels of 100, 90, 80, 70, 60, and 50%. Each confidence level is also experimented with a lower bound support of 10–100%. In both scenarios the min support of the upper bound is 100%.

Confidence level	No. of experiments	Min support lower bound	Experiment result		
			No. of association/ interrelationship rules generated	No. of cycles performed	Min support used to generate the rules
100%	5	60–100%	None		
	1	50%	7	10	50%
	1	40%	7	12	40%
	3	10–30%	10	11	35%
90%	4	70–100%	None		
	1	60%	2	8	60%
	1	50%	7	10	50%
	1	40%	7	12	40%
	3	10–30%	10	13	35%
80%	4	70–100%	None		
	1	60%	2	8	60%
	5	10–50%	10	10	50%
70%	4	70–100%	None		
	1	60%	2	8	60%
	5	10–50%	10	10	50%
60%	4	70–100%	None		
	1	60%	2	8	60%
	5	10–50%	10	10	50%
50%	4	70–100%	No rule generated		
	1	60%	2	8	60%
	5	10–50%	10	10	50%
Total exp.	60				

Table 4. Scenario and result of general association rule experiment.

5.2. Generated association rules

From the experiment, we observed that the class association mining supports the rules generated in general association mining. It also discovers interesting interrelationship (with 100% confidence level) related with the type of visit, age, altitude, temperature, and malaria type. With 100% confidence level and 60% support, outpatient cases are more closely related to the undetermined occurrence of death specifically when the age group of malaria patient is greater than 5. On the one hand, the result noted that occurrence of death is mostly related to outpatient case instead of the inpatient one. This shows that health workers offer great attention and intensive care for inpatient visits. On the other hand, most of outpatient visits are uncomplicated lab-confirmed malaria. However, the occurrence of death is undetermined and probable when the type of malaria visit is outpatient and the age of the patients is greater than 5. This may be because of lack of qualified health workers and the patients are not properly prescribed as confirmed by Ndiaye et al. [30] in Senegal that the lay health workers made negative diagnostic test.

6. Discussions

Table 5 illustrates the summary of experimental results. It is difficult to determine the occurrence of death for outpatient cases, and the experimental result revealed that the occurrence of death is related with the increment of malaria case. Roca-Feltrer et al. [31] noted that the increment of malaria cases is related with the transmission intensity, seasonality, and age that lead to a probability of occurrence of deaths. For instance, the experimental result in west Gojjam (specifically in November) supports the probability of occurrence of deaths related with the increment of malaria cases.

Knowing the seasonality of malaria helps to provide proper intervention to eradicate occurrence of death and cases [31]. Roca-Feltrer et al. [32] relate the transmission intensity, seasonality, and the age pattern of malaria and confirm that younger age groups are with increasing transmission intensity. Our experimental result also confirmed that occurrence of death is undetermined if the altitude is 1500–2000 m and when the age of the person is greater than 5 with a confidence level of 62 and 60%, respectively. This happens because of high transmission intensity. With 100% confidence level, if the type of malaria visit is outpatient and age is below 5, it is difficult to predict occurrence of deaths. Interestingly, occurrence of malaria death is related with severe anemia rather than pregnancy. As discussed by Knoblauch et al.

- If temperature is between 15 and 200°C, the type of malaria visits is outpatient, and the type of cases is uncomplicated lab-confirmed malaria, then the occurrence of death is undetermined.
- If the type of malaria is PV, then occurrence of death is undetermined, and if the type of malaria is PF, then occurrence of death is undetermined.
- If age is under 5 years and the type of malaria visits is outpatient, then occurrence of death is undetermined.

Table 5. Summarized experimental rules.

[33], anemia is prevalent in the 6- to 59-month-old children, and the association of anemia with a child age, underlying with iron requirements, is related to growth rate, and hence iron demand declines with age. Further, the algorithm associated (with 100% confidence level) with the type of malaria is unknown for inpatient malaria visits.

However, some unexpected or interesting interrelationship is prevailed such as with 100% confident level, and it is difficult to determine the occurrence of death for both PV and PF malaria types. This needs further investigation to verify whether it is unrelated or expected.

In all, the study presents the association of malaria risk factors using climate, elevation, location, type of malaria, type of malaria visits, number of cases, and death attributes. Both general and class association minings are done using Apriori techniques for discovering the association or patterns between the occurrence of deaths with the type of cases and malaria visits. And then, integrate the output of data mining (i.e., the interrelationship of risk factors) with knowledge-based reasoning. Nearest neighbor retrieval algorithms (for retrieval) and voting method (to reuse tasks) are used to design and deliver personalized knowledge-based system.

6.1. Evaluations

The evaluation of the result is executed by combining both an expert and testing tool approaches. An overall measure of pattern values, combining novelty, usefulness, and simplicity, to achieve a predefined goal is evaluated to measure the interestingness of the interrelationship. We have used different multitudes of measurement in the evaluation such as accuracy, support level, confidence, confidence level, and complexity with a 10fold cross validation. We adopted the four measures such as sensitivity, specificity, prediction accuracy, and precision to evaluate the correctness of interrelationship and validate the system through performance testing. Tenfold cross validation is used in the experiment to predict the error rate [34, 35]. The basic measure is accuracy, which computes the percentage of correctly classified instances in the test set. The accuracy of a test compares how close a new test value is to a value predicted by if-then rules [36].

The interrelationship of risk factors (association rules) was evaluated in terms of the number of rules and meaning of patterns generated at different minimum support and confidence thresholds for measuring interestingness of the rules. Association (interrelationship) was analyzed in terms of different criteria. The criteria include the number of rules generated at different minimum support and confidence thresholds. The minimum support and confidence thresholds varied from 0.1 to 1 and 0.5 to 1, respectively. As depicted in **Table 4**, at 90% confidence level with min support of 60 and 20%, the techniques generate 2 and 10 rules, respectively. Furthermore, we investigate the following indicators of the quality of the rule ranking induced by the interestingness measures of the mining algorithm in the average rank of the first rule that covers a test instance and the average rank of the first rule that covers and correctly predicts a test instance.

The performance of eHealth knowledge-based system prototype is evaluated using test case. Thus, the effectiveness of the retrieval process of eHealth knowledge-based system reasoning is measured by using recall and precision. Precision and recall are useful measures of retrieval performance [34]. Recall is the percentage of relevant cases for the query (new case) that are

retrieved, whereas precision is the percentage of retrieved cases that are relevant to the query [34, 36, 37]. Accuracy is used to measure the performance of the reuse process [34, 36].

The case similarity testing shows that when the query is made up of attribute values that have the same value with the case from the case base, the result of the global similarity becomes 1.0. But when there is a difference in the attribute values of the query and the case in the case base, the global similarity value decreases. Therefore, adding cases in the case base improves the performance of knowledge-based reasoning system in solving problems (new cases).

The nearest neighbor algorithm, which is used to develop the retrieval process of the prototype, uses distance to compute the similarity between the query and cases by representing the cases in N dimension vector. However, the recommendation doesn't have clear boundaries as it has subjectivity and depends on the experience of the domain experts as tested and adopted in [19, 23, 24]. In addition, the importance value that is assigned to the attributes of the case structure is done manually with the help of the domain experts, as there is no research that is conducted for the importance value of the attributes in malaria case management. This could affect the result of the retrieval and the reuse performance of the prototype. However, it needs user acceptance testing (using measuring usability with the system usability scale) in real-world scenarios to measure whether the potential users would like to use the proposed system frequently or not. So that, eHealth knowledge-based system for retrieving relevant cases and proposing solution will attain promising user acceptance, accuracy, and domain expert evaluation.

7. Conclusion and future work

The experimental result presents the association of risk factors (with relation to the malaria occurrence of death and type of case identification in Ethiopia) using climate, elevation, location, type of malaria, type of malaria visits, number of cases, and death attributes. Both general and class association minings are done using Apriori techniques for discovering the association or patterns of risk factors. The results noted the existence of strong association between occurrence of deaths, type of malaria visits, age, and type of cases. More interestingly, it discovers occurrence of malaria deaths, which are mostly related with severe anemia cases rather than pregnancy. It is also important to precede usability and user acceptance testing of eHealth knowledge-based system in real time and perform testing to compare and contrast with domain experts. So, health institutions have to give great attention to provide the necessary diagnosis and treatment for anemia, especially in regions that are more vulnerable for malaria. It also provides a significant contribution to design an optimal strategy in support of malaria prevention and control program within the country.

Acknowledgements

First and foremost I would like to thank God and the Holy Mother. Then, I would like to express special thanks for Worku Birhane who provided me proofreading and valuable comments.

Author details

Geletaw Sahle Tegenaw

Address all correspondence to: gelapril1985@gmail.com

Faculty of Computing, Jimma Institute of Technology, Jimma University (JU), Ethiopia

References

[1] ESTA, OSHD. Health in Africa Over the Next 50 years. 2013. www.afdb.org

[2] AfricaNext Investment Research. The future of African broadband: Economics, business models and the rise of 3G. 2010;(5)

[3] Federal Ministry of Health. Health Extension Program in Ethiopia. Addis Ababa, Ethiopia: Health Extension and Education Center; 2007

[4] Family Health Department. National Strategy for Child Survival in Ethiopia. Addis Ababa, Ethiopia: Federal Ministry of Health; 2005

[5] http://www.moh.gov.et/factsheets {Accessed: August 2017]

[6] Federal Democratic Republic of Ethiopia Ministry of Health. Ethiopia National Malaria Indicator Survey 2007 Technical Summary; 2008

[7] Ministry of Health. Malaria Prevention Control Program. 2014. http://www.moh.gov.et/malaria [Accessed: June 2014]

[8] Sahle G, Meshesha M. Uncovering knowledge that supports malaria prevention and control intervention program in Ethiopia. Electronic Journal of Health Informatics. 2014;**8**(1):e7. www.eJHI.net

[9] Sajja PS, Akerkar R. Knowledge-based systems for development. Advanced Knowledge Based System: Model, Application & Research. 2010;**1**:1-11

[10] Tan C. A prototype of knowledge-based system for fault diagnosis in automatic wire bonding machine. Turkish Journal of Engineering and Environmental Sciences. 2008;**32**:235-244

[11] Hayes-Roth F. Rule-based systems. Communications of ACM. September 1985;**28**(9):921-932. DOI: 10.1145/4284.4286

[12] Engelmore RS, Feigenbaum E. Knowledge-Based Systems in Japan. Japanese Technology Evaluation Center. Maryland: Loyola College; 1993. http://www.wtec.org/loyola/kb/toc.htm

[13] Abdelhamied K, Hafez S, Abdalla W, Hiekal H, Adel A. A rule-based expert system for rapid problem solving in crowded outpatient clinics in Egypt. IEEE. 1988;**3**:1419-1420

[14] Pandey J, Bajpai D. Developmental design of a rule based expert system for diagnosis. In: Proceedings of the First Regional Conference, IEEE Engineering in Medicine and Biology Society and 14th Conference of the Biomedical Engineering Society of India. New Delhi: An International Meet; 1995. pp. 1/41-1/42. DOI: 10.1109/RCEMBS.1995.508680

[15] Pohle C. Integrating and updating domain knowledge with knowledge discovery. In: 6th International Conference for Business Informatics 2003 (WI-2003); Dresden, Germany; September 15-17; 2003

[16] Nasiri S, Fathi M. Toward an integrated e-health based on acquired healthcare knowledge. In: 2014 Middle East Conference on Biomedical Engineering (MECBME); 2014

[17] Seera M, Lim CP. A hybrid intelligent system for medical data classification. Expert Systems with Applications. 2014;**41**:2239-2249

[18] Sedighian Z, Javanmard M. The effect of data mining on expert systems used for improving efficiency of correct speech E-learning systems. Advances in Natural and Applied Sciences. 2014;**8**(10):102-106

[19] Mulugeta T, Beshah T. Integrating Data Mining Results with the Knowledge Based System for Diagnosis and Treatment of Visceral Leishmaniasis. 2015. Available at www.ijarcsse.com

[20] Kerdprasop K, Kerdprasop N. Bridging data mining model to the automated knowledge base of biomedical informatics. International Journal of Bio-Science and Bio-Technology. 2012;**4**(1):13

[21] Martin B. Instance-based learning: Nearest neighbour with generalisation [MSc thesis]. University of Waikato; 1995. Available at: http://www.cs.waikato.ac.nz/pubs/wp/1995/uow-cs-wp-1995-18.pdf [Accessed: March 27, 2011]

[22] Mishra D, Sahu B. Feature selection for cancer classification: A signal-to-noise ratio approach. International Journal of Scientific & Engineering Research. Hamilton, New Zealand: University of Waikato. 2011;**2**(4):1-7

[23] Salem AM. Case-based reasoning technology for medical diagnosis. World Academy of Science, Engineering and Technology. 2007;**31**:9-13

[24] Bekele H. A case based reasoning knowledge based system for hypertension management [MSc thesis]. Ethiopia: Addis Ababa University; 2011

[25] Han J, Micheline K. Data Mining: Concepts and Techniques. Waltham, USA: Morgan Kufman Publishers is an imprint of Elsevier; 2006

[26] Larose DT. Discovering Knowledge in Data: An Introduction to Data Mining. New Jersey, Canada: John Wiley & Sons; 2005

[27] Witten IH, Frank E. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Series in Data Management Systems. 2nd ed. Physical Description xxxi. San Francisco: Morgan Kaufman Publishers; 2005. 525 p. ill. https://trove.nla.gov.au/version/46617902

[28] Wu X, Kumar V, Ross Quinlan J, et al. Knowledge and Information Systems. Springer-Verlag. 2008;**14**:1. https://doi.org/10.1007/s10115-007-0114-2. Print ISSN: 0219-1377. Online ISSN: 0219-3116

[29] The University of Waikato, WEKA Manual (Waikato Environment for Knowledge Analysis) for Version 3-7-4, This manual is licensed under the GNU General Public License version 2. Available at http://www.gnu.org/copyleft/gpl.html

[30] Ndiaye Y, Ndiaye JLA, Cisse B, Blanas D, Bassene J, Manga IA, Ndiath M, Faye SL, Bocoum M, Ndiaye M, Thior PM, Sene D, Milligan P, Gaye O, Schellenberg D. Community case management in malaria: Review and perspectives after four years of operational experience in Saraya district, south-East Senegal. Malaria Journal. 2013;**12**:240

[31] Roca-Feltrer A, Armstrong Schellenberg J, Smith L, Carneiro I. A simple method for defining malaria seasonality. Malaria Journal. 2009;**8**:276

[32] Roca-Feltrer A, Carneiro I, Smith L, Schellenberg JRMA, Greenwood B, Schellenberg D. The age patterns of severe malaria syndromes in sub-Saharan Africa across a range of transmission intensities and seasonality settings. Malaria Journal. 2010;**9**:282

[33] Knoblauch AM, Winkler MS, Archer C, Divall MJ, Owuor M, Yapo RM, Yao PA, Utzinger J. The epidemiology of malaria and anaemia in the Bonikro mining area, central Côte d'Ivoire. Malaria Journal. 2014;**13**:194

[34] McSherry D. Precision and recall in interactive case-based reasoning. In case-based reasoning research and development (ICCBR). Lecture Notes in Artificial Intelligence. 2001;**2080**:392-406

[35] Kohavi R. A study of cross validation and bootstrap for accuracy estimation and model selection. In: The International Joint Conference on Artificial Intelligence; 1995

[36] Junker M, Hoch R, Dengel A. On the evaluation of document analysis components by recall, precision, and accuracy. In: Proceedings of the Fifth International Conference on Document Analysis and Recognition; 1999. pp. 713-716

[37] Losee RM. When information retrieval measures agree about the relative quality of document rankings. Journal of the American Society for Information Science. 2000;**51**: 834-840

Edited by Thomas F. Heston

eHealth has revolutionized health care and the practice of medicine. Internet technologies have given the most rural communities access to healthcare services, and automated computer algorithms are improving medical diagnoses and speeding up the delivery of care. Handheld apps, wearable devices, and artificial intelligence lead the way, creating a global healthcare solution that is smarter and more accessible. Read what leaders in the field are doing to advance the use of electronic technology to improve global health.

Published in London, UK

IntechOpen

ISBN 978-1-83881-621-6