IntechOpen

Quality Control and Assurance

An Ancient Greek Term Re-Mastered

Edited by Leo D. Kounis

QUALITY CONTROL AND ASSURANCE - AN ANCIENT GREEK TERM RE-MASTERED

Edited by **Leo D. Kounis**

Quality Control and Assurance - An Ancient Greek Term Re-Mastered
http://dx.doi.org/10.5772/63135
Edited by Leo D. Kounis

Contributors

Ching-Chow Yang, Dietmar Winkler, Marta Sabou, Stefan Biffl, Beata Szetela, Agustín G. Asuero, Julia Martín, Nieves María Velázquez, José González, F.J. Domínguez-Mayo, J.A. García-García, M.J. Escalona, Rodrigo De Santis Neves, Daniel Pereira Da Silva Fernandes, Carlos Eduardo Cardoso Galhardo, Erlon H. Martins Ferreira, Rafael Mello Trommer, Jailton Damasceno, Tyler Florio, Joseph Valadez, Lusine Mirzoyan, M. Mejías

Notice

First published in Croatia, 2017 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019.
IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

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

Quality Control and Assurance - An Ancient Greek Term Re-Mastered
Edited by Leo D. Kounis

p. cm.

Print ISBN 978-953-51-2921-9

Online ISBN 978-953-51-2922-6

eBook (PDF) ISBN 978-953-51-6696-2

Meet the editor

Leo D. Kounis is a Quality and Manufacturing Systems Engineer and Departmental Manager at the Manufacturing Directorate of the Hellenic Air Force's State Aircraft Factory, KEA, and is acting as Scientific Advisor to the Hellenic National Defence College.
He obtained his BEng (Hons) degree in Manufacturing Systems Engineering, his MSc degree in Quality Engineering and his PhD degree in Systems Reliability from the University of Hertfordshire, United Kingdom. Dr. Kounis has worked as a Senior Quality Engineer in a number of private companies in Greece and has acted as a part-time lecturer at a number of universities.
His research interests focus on the area of quality, transportation and sustainable energy. He has published a number of scientific papers.

Contents

Preface

This book on *Quality Control and Assurance* showcases the evolution of quality concepts, techniques and tools that are implemented by manufacturing companies and service providers.

The term *quality* stems from the Latin *qualitas*, which was introduced by Marcus Tullius Cicero, who created a Latin philosophical vocabulary, combining the Platonic method and the Roman philosophy. His influence to Latin and European languages is truly immense. Cicero was inspired by Plato's Academy, and it is indeed Plato and Aristotle who use the term *poiotes* originating from *poios.*

Aristotle, in his work *Categories,* discusses various forms of *qualities.* It is worth noting at this point that *quality* was defined by philosophers as an attribute, or characteristic of an object, so as to distinguish between certain levels and kinds of associated (given) or observed qualities.

Scientific findings show that Ancient Greeks did utilize qualitative control and non-destructive testing (NDT) so as to pursue *good* quality. This included—but was not limited to—the manufacture of bronze fittings (concerning the erection of columns, such as the Philonian Stoa), to quality control of gold and silver coins.

The term *quality* begins to evolve in the science that is known today, when Walter Shewhart developed—in 1924—the notorious control chart. W. E. Deming, J. Juran and A. Feigenbaum, among others, further elaborated on the work of W. Shewhart and set the foundations of *total quality management (TQM).*

Japan is regarded as the forerunner to establishing the framework of company-wide quality control, so as to increase its market share in the West. Scholars and business insiders alike acknowledge its influence. In order to counterbalance its momentum, the British Standard BS 5750 was developed, aiming at improving quality practices in an increasingly antagonistic environment. The advent of the International Standards Organization (ISO) 9000 became the internationally recognized quality management system standard. So much so that *Chrysler, Ford and General Motors*—although possessing their own quality manuals—oriented their efforts in accordance to the ISO 9000 series of standards, which formed the ISO/TS 16949 series of quality standards.

The *Malcolm Baldrige National Quality Award (MBNQA)* is only a handful of awards that promotes company excellence and successful performance. It goes without saying that the word *excellence* was soon adopted by enterprises and state legal entities, such as the *European Foundation for Quality Management (EFQM)* and the *British Quality Foundation (BQF)*, to enhance qualitative practices. As such, it can be said that quality forms a holistic model that finds widespread application.

This book comprises nine chapters that cover a diverse area and is split in two sections.

The first section may be viewed as an introductory part that elaborates on the evolution of quality concepts and assurance by using statistical tools and techniques. In doing so, it presents a model that aims to capture customer's emotional responses in opting for a specific product or service, which, in turn, may increase companies' market influence.

The ISO/IEC 17025 is the main ISO standard used by testing and calibration laboratories and is as such more explicit regarding competence-specific requirements. The presented paper discusses key objects and strategies for laboratories to effectively overcome potential bottlenecks.

Youden plots are a simple graphical technique to analyze inter-laboratory results and acknowledge variations. The next study illustrates a method that is applied to determine contaminants in surface waters and elaborates on findings in chemistry.

Lot quality assurance sampling (LQAS) is a random sampling technique first developed in 1920, aimed for use in manufacturing. However, LQAS has become an accepted sampling method in the area of public health that is recommended by the World Health Organization (WHO). Among others, it is used to monitor the prevalence of abortions and the quality of reproductive health care in Armenia.

The second section discusses quality frameworks and management in a number of enterprises. TQM, although providing a platform for companies to improve their performance, was focused on the private sector and did not acknowledge the management of the life cycle of large infrastructure portfolios.

In order to bridge the gap, the Andalusian Public Administration in Spain called for the THOT project, so as to study in detail enterprise content management systems (ECMS). In doing so, a quality evaluation framework (QuEF) is proposed, intended to assist organizations in selecting the most suitable ECM system.

The relationship between supply chain risk management (SCRM) and quality management is assessed in the next paper, and a link between TQM and SCRM schemes is identified.

The advent of information technology (IT) and the complexity of modern enterprises operating in a globalized economy necessitated the interaction between technology and quality. This resulted in the development of a number of quality management software packages, such as enterprise resource planning (ERP), supply chain management (SCM) and customer relationship management (CRM), to name a few. Enterprise quality management software (EQMS) is a software platform that establishes cross-functional communication whilst standardizing, centralizing and streamlining management-specific data from across the value chain. Preconfigured and industry-specific IT solutions are available that differ in design, functionality and capabilities.

To this end, the ISO/IEC/IEEE 29119 software testing standards intend to support a variety of software-specific life cycles and methods. The process per se is risk-based testing. It facilitates the development of a strategy and supports test planning.

Indeed, model-driven engineering (MDE) is yet another software tool that concentrates on further developing conceptual models and is based on updated company-specific best practices. It aims at optimizing in-between system compatibility whilst facilitating communication between managers, designers and users of the application. In this context, established quality assurance (QA) methods need to be adapted to MDE so as to eliminate—among others—potential bottlenecks between incompatible data models.

In view of the above, the ALAMEDA software package provides support to life cycles focused on the generation, implementation and testing from its initial phases. In doing so, it establishes a degree of compliance on par with the ISO/IEC-29119 standard. The outcome thereof—in conjunction with the implication of MDE—shows an impact in reducing time-to-market and improved software-specific quality.

Semantic Web technologies (SWT) promote common data formats and exchange protocols on the web. They aim at bridging the gap between heterogeneous data models and sources and set the platform for establishing high levels of quality assurance. As such, the next paper introduces a set of success-critical requirements aimed at efficiently detecting shortcomings in MDE-governed environments whilst proposing a defect detection framework (DDF).

Virtual currencies are digital money that is not regulated and is controlled by its developers and circulated among users of a virtual community. The Bitcoin is the forerunner thereof and gained in popularity. The final chapter researches its volatility by comparing it to the US dollar (USD) and the Euro. This is done by means of control charts.

This book features a number of selected topics showcasing current implications of quality control and assurance.

It is hoped that it will serve as the basis for further research.

Dr. Leo D. Kounis
Hellenic Air Force
State Aircraft Factory, KEA,
Greece

Section 1

The Evolution of Quality and the Application of Quality Technique Tools

Chapter 1

The Evolution of Quality Concepts and the Related Quality Management

Ching-Chow Yang

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/67211

Abstract

Enterprises usually adopt some quality practices to control the product quality during the manufacturing process in order to assure the delivery of qualitative good products to customers. The quality practices or quality management systems adopted by industries will further evolve due to the changes of quality concepts as time goes by. This chapter discusses the change of quality concepts and the related revolution of quality management systems in the past century. The quality concepts were gradually changed from the achievement of quality standards, satisfaction of customer needs, and expectations to customer delight. Since merely satisfying customers is not enough to ensure customer loyalty, the enterprises gradually focus on customers' emotional responses and their delight in order to pursue their loyalty. The emotion of "delight" is composed of "joy" and "surprise," which can be achieved as the customers' latent requirements are satisfied. Thus, the concept of "customer delight" and the means to provide the innovative quality so as to meet the unsatisfied customers' latent needs are elaborated on. Finally, a framework of innovation creation is developed that is based on the mining of customer's latent requirements. This outline will manifest the essential elements of the related operation steps.

Keywords: quality concept, customer satisfaction, customer delight, customers' latent requirement, innovation creation system

1. Introduction

It is widely recognized that consumers only buy the goods with good quality, desired functions, and accepted price [1]. Therefore, industries always adopt several management practices, or even develop some management systems to design and produce the products so as to meet customers' needs and expectations showcasing good quality and lower costs. But the

management practices or systems adopted by the industries are always changed due to the changes of the quality concepts as time goes by. This occurs as several quality gurus give the pragmatic definitions of "quality" over different time periods, thus causing industries to implement different practices or systems to control overall product quality [2].

In the first three decades of the last century, quality was defined as "conforming to the standards and specifications of a product" [3]. Thus, the commonly adopted quality practices by industries were the standardization of quality, inspection, and rework. Deming emphasized that "quality is to fulfill the requirements of customers and satisfy them" [4]. Hence, the meaning of quality was gradually changed to a "customer-focused" perspective. Enterprises, therefore, committed themselves to satisfy customers' needs and expectations. Their aim was to pursue customer's satisfaction and loyalty [5, 6]. Companies also developed a number of methods to find out customers' needs and expectations. To this end, in-depth interviews with customers were performed, customer surveys conducted, and market research was done. But, when Apple announced several innovative products, and their sales were increasing, it became apparent that only satisfying customers' requirements are not enough [7]. As a matter thereof the identification and fulfillment of customers' unsatisfied latent needs was gauged in conjunction with their emotional responses [8].

Entering the new century, several studies have suggested that merely satisfying customers is not enough to ensure their loyalty [9–11]. The industries need to focus on the customer's emotional responses and provide the products with attractive quality in order to pursue customers' delight [11]. It is worth mentioning at this point that Apple is regarded as a future trend setter, as it successfully created several innovative products such as iPod, iPhone, and iPad, triggering increased sales. A strong customer service department equally assisted in causing customers' delight experiences [12].

In view of the above, quality concepts have changed. Terms such as "customer delight" are deemed as forming the crucial elements of quality concepts, which are coined as "attractive quality" and "innovative quality" [13]. The conceptualization of delight is that the emotion of "delight" is composed of "joy" and "surprise," which can be achieved as the customers' latent requirements are satisfied. The provision of innovative quality products is the strategic tool to meet the unsatisfied customers' latent needs and their curiosity. These changes of quality concepts will lead the enterprises to reengineer their existing quality system, in order to develop the innovative quality attributes of products and services alike so as to retain and attract customers. This scenario causes the quality professionals and researchers to further develop and expand the new quality system beyond TQM. Based on performed research it is the author's view that research should focus on developing an effective new quality system.

During the past century, several quality gurus had led the changes of quality concepts. This chapter will discuss the changes of quality concepts during different time periods over the past century. To this end, it will also state the reflected revolutions of quality management systems with respect to the changing quality concepts. These are arranged in Sections 2 and 3, respectively. The development of "total quality management" (TQM) philosophy in the 1980s was an important landmark, which caused the change of new quality concepts and the reengineering of the quality management system, as introduced in Section 4. Based on

http://dx.doi.org/10.5772/67211

the performance evaluation of the TQM implementation of TQM, it is worth noting at this point that the "business excellence model" had been proposed during the 1990s. This is elaborated in Section 5. This chapter also introduces the development of an integral model of a business excellence system beyond TQM, based on the realization of TQM.

Section 6 expands on the changes of new quality concepts pertaining to "customer delight" and "innovative quality" based on the investigation of customers' latent needs in the new century. This chapter includes a system, which may assist and cater for identifying and foster upon customers' undersatisfied latent needs and delights, enhancing innovative quality, which is addressed in Section 7. The conclusion of this chapter is listed in the final section.

2. The early quality concepts and the reflected quality management system

Since the early twentieth century, manufacturing processes and activities initiated the control practices to assure the product quality [14]. Manufacturing companies focused on the related productivity and manufacturing costs. As such, the quality concept and control were product-focused.

2.1. The quality concept of "standard" and the "inspection" control

In 1913 Ford Motor Company, FMC, created the assembly line in their newly opened factory in Highland Park, Michigan due to the influence of the scientific management of Frederick W. Taylor. This resulted in Ford increasing its manufacturing volume [15]. Ford's assembly line was copied by many manufacturing companies. However, companies turned their attention to control product-related quality issues. Since manufacturers were more "product-focused" in that time, the quality concept was, therefore, aimed at "conforming to the standards and specifications of a product" [3]. This in turn, impelled quality engineers in manufacturing industries to implement the method of "inspection" so as to control the quality of a manufactured product.

Product designers and process engineers designed the standards and specifications of the products, which were based on their critical attributes. They also set up the standards of the manufacturing process and the standards of operations (SOPs). In doing so, the involved workforce was requested to perform the tasks according to the developed SOPs. The quality inspectors checked the dimensions and characteristics of products, detected the errors and failures, and took the necessary actions to improve the product quality.

2.2. The development of process quality control

In order to ensure product quality, companies needed to utilize the "full inspection" method. This was costly, since it required much time and labor efforts, and resulted in high internal quality costs [16]. Walter Shewhart, thus, created the control chart, a quality technique tool that he pioneered in Bell Laboratories working as a quality control engineer [17]. He proposed the use of a sampling inspection method instead of a 100% inspection to reduce the overall

amount of inspection. The control chart was utilized to monitor the quality performance concerning the critical aspects of a process, whereas the attributes of the product were identified by means of sampling methods [18]. This enhanced overall effectiveness and also reduced the associated costs.

Sampling inspection and control charts use many statistic tools such as probability theory, the methods of random sample, analyses of sample mean and deviations to name but a few, as means for improving quality levels. Hence, the method of quality control suggested by Shewhart was called "statistical process control" (SPC) or "statistical quality control" (SQC). Sampling inspection may not always ensure product quality, as it might cause fewer defective products to be shipped to customers [19]. As a result, thereof, it leads to extra outside quality costs. Shewhart, however, had argued that if the missed number of defects was small, then the savings in inspection costs made it worthwhile [20].

3. The quality concepts and the reflected quality management system in the mid-twentieth century

More often than not, quality and the associated price of the goods are primary factors that are considered by customers prior to them materializing a purchase [21]. As such, manufacturing companies mainly emphasis on the control of quality and costs during the manufacturing process, especially the ones that trigger poor quality.

3.1. The development of "quality costs"

In essence, Juran propounded the concept of "quality costs" in his book "*Quality Control Handbook*" in 1951 [22]. Juran subdivided the quality costs into prevention, appraisal, internal failure, and external failure costs. The performed literature review indicates that the losses due to the production of defects and manufacturing failures are more than the costs of quality control; in particular, the costs caused by internal and external failures. In view of the above, the implementation of SPC could not effectively control the quality costs.

Despite this general classification being still widely used by practitioners and researchers, several authors had suggested other kinds of quality costs, especially the "invisible" or "hidden" costs. The term hidden cost is used to indicate failure costs that are inadequately recorded in company accounts and/or failure costs that are never actually discovered. Yang addressed the hidden costs through the definition and addition of two new categories: "extra resultant cost" and "estimated hidden cost" [16]. As such, the quality costs can be subdivided into the following six categories:

- prevention costs
- appraisal costs
- internal failure costs

http://dx.doi.org/10.5772/67211

- external failure costs
- extra resultant costs and
- estimated hidden costs

Figure 1 depicts the new definition of quality costs.

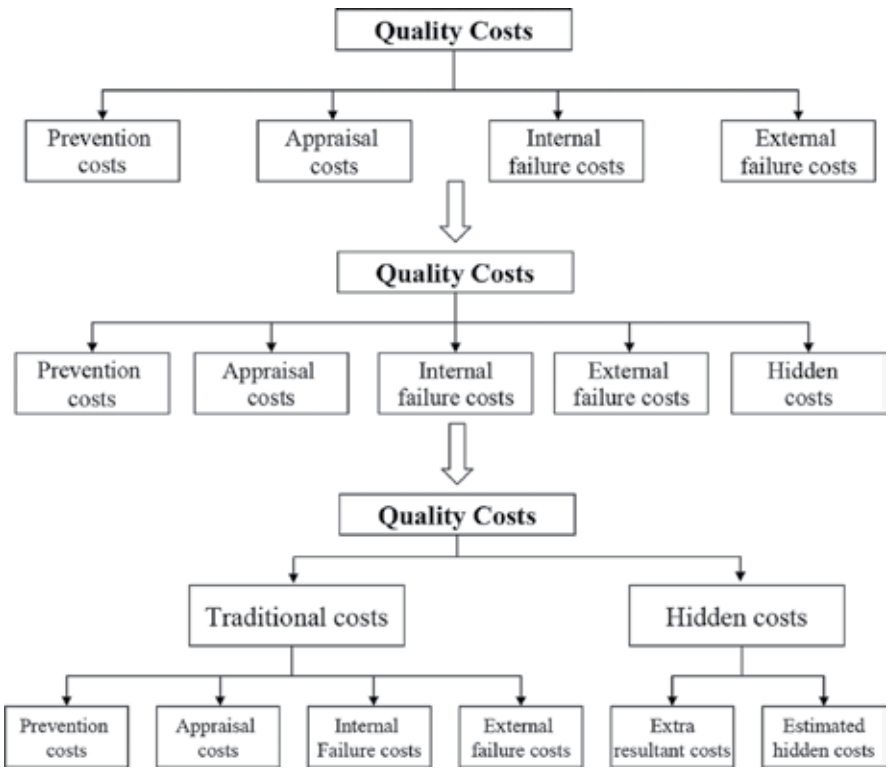

Figure 1. Definition of categories of quality costs.

3.2. The quality concept "quality assurance" and the related TQC-system

After Juran and several quality experts emphasized the quality cost issue, the concept of quality costs was widely accepted by industry. Meanwhile, another quality concept appeared gradually, namely "quality assurance." However, a large number of experts in the area of quality control, including Feigenbaum, asserted that implementing an SPC system on its own could not effectively control quality costs [23]. It is worth mentioning that the concept of "quality assurance" was "users-oriented," implying that "product possesses the fitness for purpose of use based on its functions" and hence "quality is zero defects and meeting the specifications 100%" [24].

The performed literature review suggests that Feigenbaum emphasized that quality assurance could not be achieved by placing the control just on production processes. Thus, the concept of total quality control, TQC, was introduced in 1959. TQC emphasizes that product quality needs to be implemented at all stages of the product life cycle. The sequence of quality activities is briefly shown as:

- product design,
- incoming quality approval,
- process quality control,
- product reliability,
- inventory,
- delivery, and
- customer service.

Actually, Feigenbaum's quality concept and ideas were similar to those described by Deming, Juran, and Crosby [25].

The concepts and approaches of SPC, TQC, and "costs of quality" were introduced in Japan during 1960 by Deming and Juran. The Union of Japanese Scientists and Engineers (JUSE), which was formed in 1946, synthesized the concepts, principles, and approaches of statistical process control and total quality control [26]. JUSE promoted the practices of TQC and the quality concepts pursuing the zero defect culture and executing the task right first time.

4. The age of total quality management

While Japanese industries adopted the TQC practices, they emphasized the education and training of quality for all employees and the cultivation of a quality culture. Therefore, the implementation of TQC in Japanese industries was very different from the original TQC.

4.1. The emerging of the Japanese company-wide quality control, CWQC

Actually, the Japanese TQC possessed several critical characteristics listed as follows:

- customer-focused and quality-first as the quality policies,
- full participation and teamwork,
- education and training of quality for all employees,
- realization of "do the right thing first time,"

http://dx.doi.org/10.5772/67211

- concept and materialization of a "zero defect" culture,
- "continuous improvement" as the key quality activity,
- everyone is responsible for achieving high quality levels,
- emphasizing on the prevention activities and quality assurance,
- cultivating a quality culture environment.

Based on aforementioned critical characteristics, the Japanese TQC was acknowledged as company-wide quality control (CWQC).

The implication of CWQC in conjunction with Japanese industrial competitiveness and strategic advantages facilitated their entrance into western markets. Japanese enterprises enjoyed an increase in global market share by providing the customers with high-quality products at lower prices [27]. This in turn, resulted in western companies facing an increased amount of competition from Japanese and other Asian manufacturers.

4.2. The development of total quality management

The Japanese competition caused American and western industries to implement benchmarking projections by studying Japanese CWQC performances. In doing so, they adopted the principles of Japanese quality management. Based on the critical characteristics listed above, western professionals further developed and refined CWQC to become a total quality management (TQM) system. Its fundamental principles are listed as follows:

- "customer-focused" management,
- "continuous improvement" as the key quality activity,
- top management's promise persistence for pursuing quality,
- full participation and teamwork,
- education and training of quality for employees,
- employees' good quality concept,
- quality leadership,
- long-term supplier relationship,
- implementation of quality management system,
- cultivation of quality culture.

It is the author's view that the TQM was well developed and suitable for western organizations, as depicted in **Figure 2**. Thus, it was widely adopted by industries and nonprofit organizations around the world.

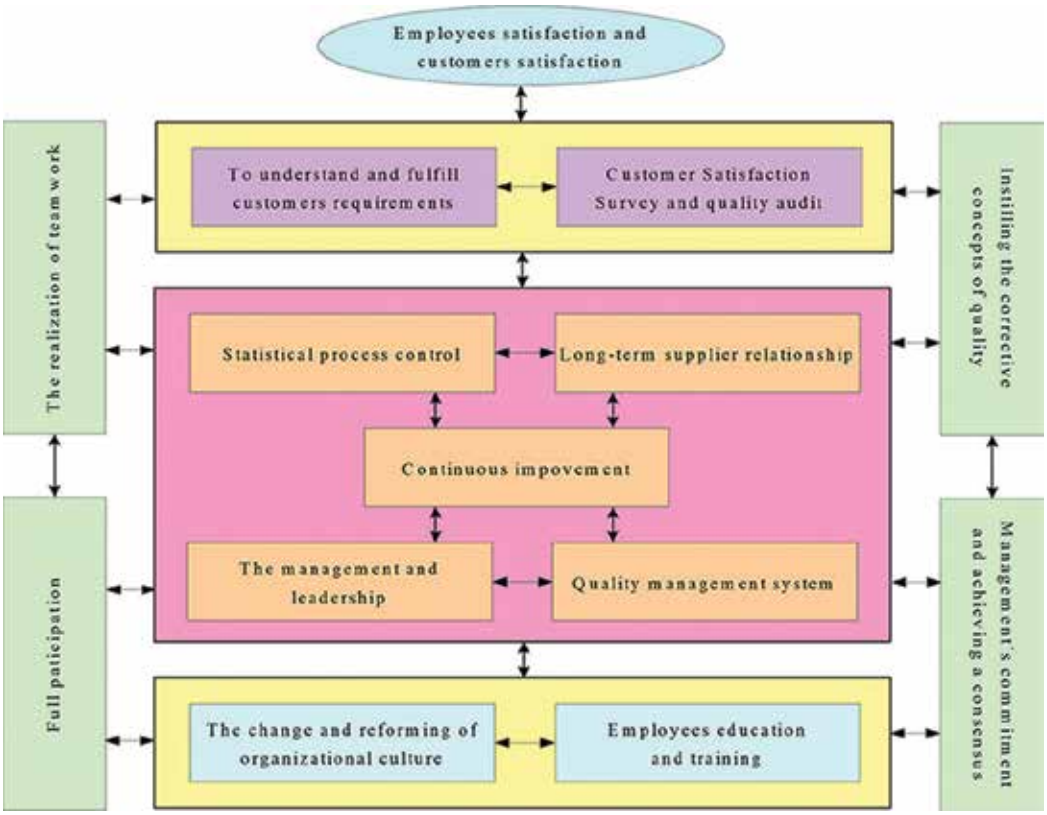

Figure 2. The framework of TQM.

The development of TQM was also influenced by the western quality experts, namely Deming, Juran, and Crosby [28]. As already stated Deming's quality concept was that "quality is to fulfill the requirements of customers and satisfy them" [4]. Crosby's definition of quality was also similar, as he defined quality as "conformance to customers' requirements." Juran also defined quality as being "fitness for purpose of use, …, it is judged by the users, not the manufacturers, or the merchants" [22]. TQM was thus an integrated model of management philosophies, quality concepts, and a set of practices, which were also influenced by Deming's 14 points and Juran's quality trilogy. However, to implement the TQM successfully it is necessary to integrate the so-called "hard side" elements (that is, statistical methods, quality control tools, process standardization, and improvement, etc.) with the "soft side" aspects (that is, quality concept, employees' participation, education and training, and quality culture, etc.) [29].

From the mid-1980s onward, several important quality programs were being launched. Besides the development of TQM, the ISO 9000 system and the Six-Sigma program (which was initiated by Motorola) commenced in 1987. It is worth noting that up to date, the ISO system has had four revisions in 1994, 2000, 2008, and 2015, respectively.

http://dx.doi.org/10.5772/67211

The Six-Sigma quality scheme was being widely imitated by GE in 1995 [30]. The successful implementation of Six-Sigma by Motorola, GE, and several other multinational companies caused the Six-Sigma philosophy to be globally adopted by a number of industries and organizations [31].

5. From Malcolm Baldrige National Quality Award to the Business Excellence Model

From the late 1990s to present, the traditional quality concept has seen a number of changes. In the United States of America, the Department of Commerce introduced in 1987 the Malcolm Baldrige National Quality Award (MBNQA) by benchmarking the Japanese Deming Award [32] to promote the implementation of TQM in industries and nonprofit organizations.

5.1. The national quality award and TQM

The main aim of TQM implementation is to achieve customers' satisfaction [33]. This is turn results in a company improving its financial performances [34–36]. The benefits may be seen in areas such as cost reduction, increased market share and profit, and enhanced business competitiveness [37]. As a result, MBNQA specifically emphasizes business excellence. Besides the implementation of TQM, MBNQA also considers the strategic management, and information management and analyses.

The European Foundation for Quality Management (EFQA) also launched in 1992 a European Quality Award (EQA) [38]. Based on the performed research, more than 80 countries have established National Quality Awards (NQAs). These NQAs are largely based on the MBNQA, EQA, and the Deming Prize [39, 40], and are generally considered to be an effective way of pursuing business excellence, including customer loyalty and better long-term financial performances. Most of these national quality awards may form part of the TQM system namely, strategic management, performance evaluation, human resource management, and even IT system and innovation.

5.2. Prospects of the business excellence model beyond TQM

The framework of the National Quality Award is indeed beyond the extent of TQM. In most of the cases, National Quality Award such as MBNQA and EQA are acknowledged as a business excellence model, with the industries adopting it as their strategic business system. Yang [41] proposed a framework of an integral model of a business-excellence system based on the integration of TQM, MBNQA, and EQA, and its islands:

- strategic management system,
- performance evaluation, and
- innovation (**Figure 3**).

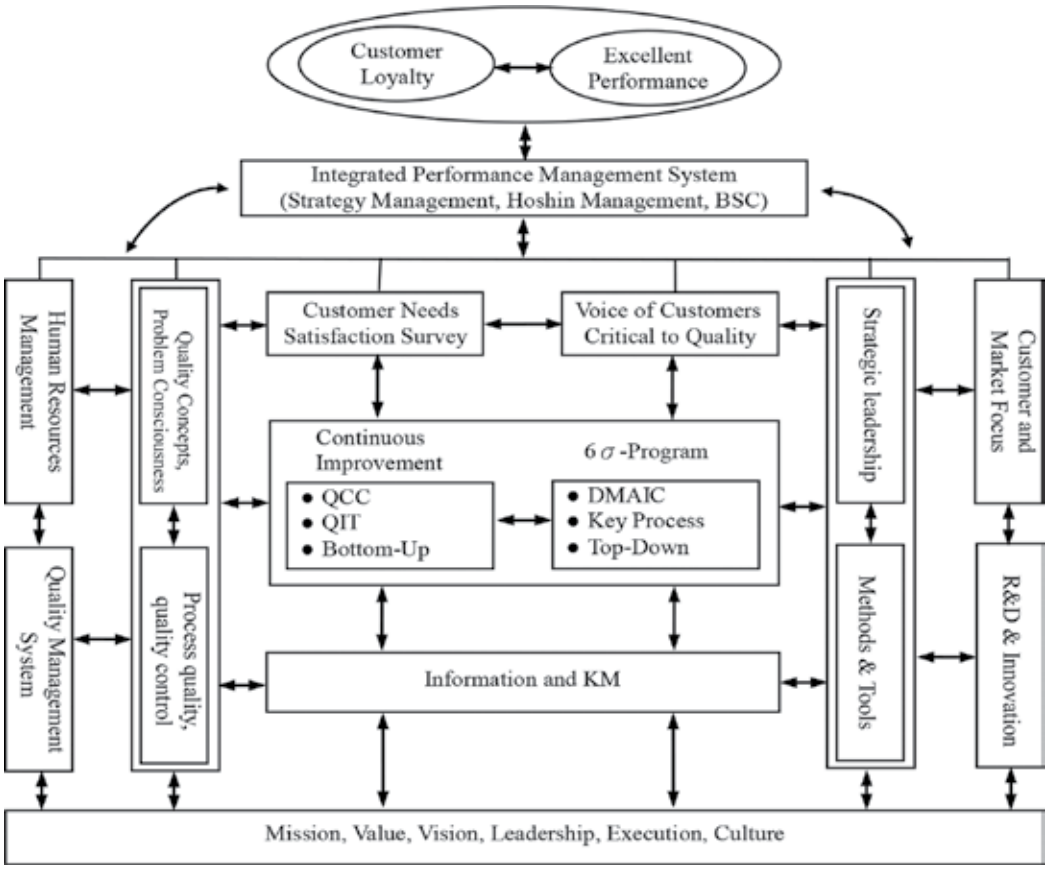

Figure 3. The framework of the proposed integral model of a business-excellence system.

Meanwhile, the concept of "sustainable development" was also widely adopted in industry [42, 43]. Zairi asserted that "sustainable development is based on a perceived need to address environmental deterioration and to maintain the vital functions of natural systems for the well-being of present and future generations" [44]. He thus, proposed an integrated model that incorporates the sustainable development into a TQM system. Yang et al. developed a framework of the system of environmental management, which constitutes of 30 activity items themselves belonging to 12 initiatives [45].

6. The concepts of "customer delight" and "innovative quality" in the new century

A critical strategy of pursuing business excellence is to raise customers' loyalty, which cannot be assured by grossly satisfying customers. Indeed, even satisfied customers would defect at

http://dx.doi.org/10.5772/67211

a significant rate in many industries. It is recognized that customers' satisfaction is the result of the fulfillment of their explicit needs. As such, satisfaction only of their needs does not directly result in customer loyalty, since the competitors would also fulfill these explicit needs with better customer-valued products [46].

Several studies have revealed that the degree of loyalty depends on whether the customers are "totally satisfied" or merely "satisfied" [47, 48]. The totally satisfied customers are actually about five times more likely to repurchase a product or service than those who are merely "satisfied" [49, 50]. This suggests that enterprises must strive for "total customer satisfaction," or even "customer's delight" [51].

"Customer delight" is a new quality concept. A conceptualization of "delight" is the customer's emotional response, which is composed of "joy" and "surprise"—both encountered in the providing process of goods/services [52]. In order to delight customers, the enterprises need to provide the goods/services with attractive and innovative quality incentives. The strategic actions are to identify customers' latent needs and to create customer value by developing the innovative products and attractive services in order to fulfill these.

If a company brings out some innovative products that consumers did not know they needed before, then these innovative products will often spark high demand [53]. Indeed, Moore illustrated eight types of innovations, namely [54]:

- application innovation,
- product innovation,
- process innovation,
- experience innovation,
- marketing innovation,
- business model innovation,
- structural innovation, and
- disruptive innovation.

These kinds of innovations, especially the first four types, will result in significant effects on the fulfilment of the customers' latent needs and their delight experiences.

7. The development of the innovation system

In order to raise the customers' loyalty, the enterprises should develop the core capabilities with an innovation system. In view of the above, the prerequisite for developing such a system is been described. As such, this chapter proposes a systematic innovation development system based on the product/service value chain.

7.1. The framework of the innovation system

At first, the differences among improvement, reengineering, and innovation are established as shown in **Figure 4**. Improvement is implemented on the existing processes, as some problems may have been encountered during the manufacturing process or the service delivery process. Continuous improvement is the key activity of the TQM system, its aim is to improve the quality of products/services and all the observed and or experienced bottlenecks on the processes. The final objective is to raise customer satisfaction [55].

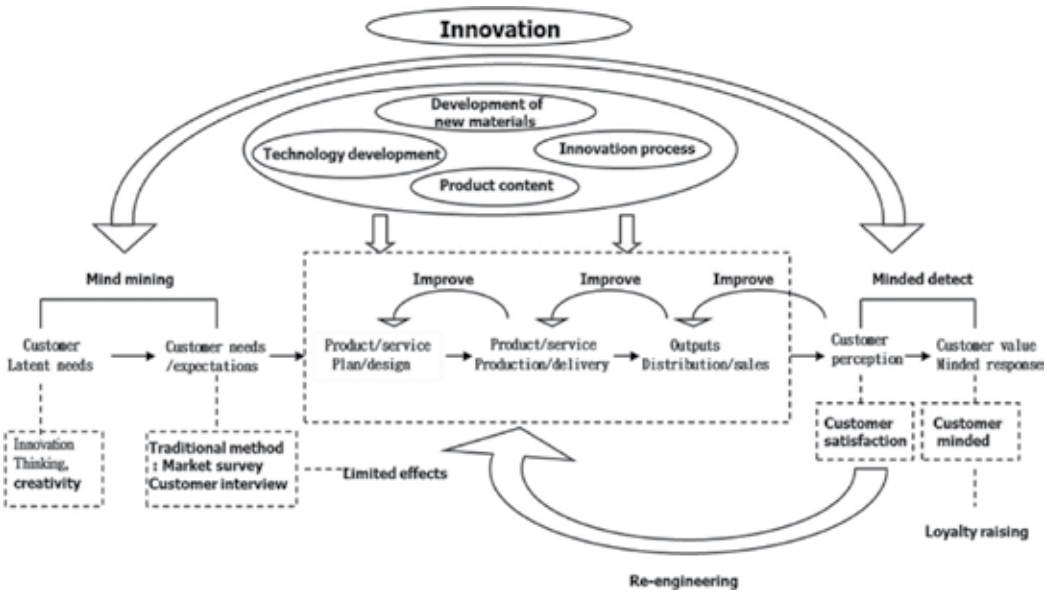

Figure 4. The framework of the innovation system.

Hammer and Champy asserted that improving solely the production system or service delivery processes may not result in significant business performance [56]. Sometimes it is needed to reengineer the critical processes, or to redevelop the provision systems based on customer voices. The latter applies in particular to key customer values which are difficult to fulfill by adopting the improvement actions, as shown in **Figure 4**. Thus, the aim of reengineering is to create customers' values, which are the critical factors of raising customer loyalty. Besides, reengineering also reaps the benefits including shorter delivery times of products/services (including lead time and production time), reduction of costs, and the effective utilization of resources.

Most companies develop and implement enterprise resources planning (ERP), customer relationship management (CRM), and supply chain management (SCM) simultaneously [57]. By integrating these systems, enterprises can exert excellent business' performance and customer's loyalty. In order to effectively utilize these networks and their integrated systems, it is the author's view that companies should proceed with the necessary processes of reengineering. This in turn, may have an impact on the critical functional activities on a timely basis.

http://dx.doi.org/10.5772/67211

The utilization of high-tech functions and the rapid application of the internet resulted in the change of customers' purchasing behavior [58, 59]. Customers now focus not only on the evaluations of price and functions; rather they also integrate quality with the perceived value [60]. It is the author's view that business systems ought to change and pursue innovation, speed, and quality if they want to satisfy the overall needs of customers and create new value.

Customer satisfaction may be termed as fulfilling their requirements and expectations. This, however, is not enough for delighting them and raising their loyalty. Performed research indicates that a number of customer's latent needs are never discovered and satisfied by enterprises [61]. Thus, the latter's competitive force is to probe the unmet customers' latent needs. There are some new methodologies proposed by several researchers to pinpoint customers' latent needs [62]. Once the latter are identified, companies ought to overcome bottlenecks and aim at developing and providing innovative products/services to satisfy customers' latent needs and curiosity.

Eventually the customer will be set at the center of innovation. Thomke and von Hippel asserted that "tapping into customer innovation can certainly generate tremendous value, but capturing customer value is hardly a simple or straightforward work" [63]. The term "innovation" is usually considered as a process or a capability to create new ideas, and then develop the innovative products/services in order to meet customers' unsatisfied latent needs [64, 65]. In essence, the realization of "innovation" depends on the implementation of an innovation system, as shown in **Figure 4**. This innovation system starts at the "mind" mining of the customer's latent needs and ends in the customer's value and response, and vice versa.

7.2. The transfer loop of innovation system

In order to effectively implement the innovation system, a transfer loop of the innovation system has been developed, as shown in **Figure 5**. It constitutes of four constructs:

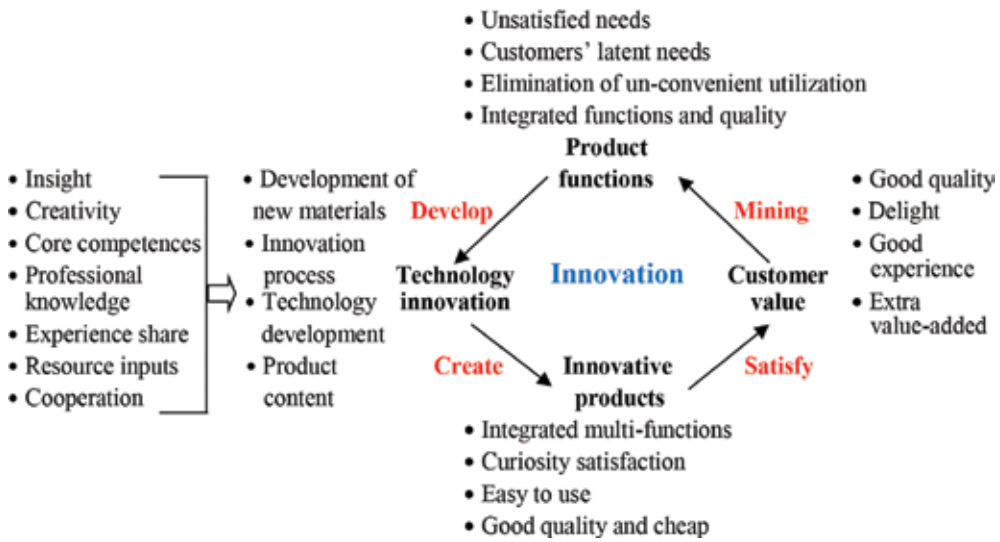

Figure 5. The transfer loop of the innovation system.

- technology innovation,
- innovative products,
- customer value, and
- product functions.

It is worth mentioning that each construct contains several items. As such, the construct "technology innovation" is termed as the fundamental force of innovation, which contains the following items:

- technology development,
- development of new material,
- product content, and
- innovation process.

These are driven by the following core activities:

- insight,
- creativity,
- core capability,
- professional knowledge,
- experience share,
- resource inputs, and
- cooperation.

The effective implementation of "technology innovation" can create the "innovative products," which includes the following:

- integrated multifunctions,
- curiosity satisfaction,
- easy to use, and
- good quality at an affordable price.

Each of the other three constructs also contains several items. These constructs and their involving items are shown in **Figure 5**.

The features of "innovative products" will satisfy the "customer value" which is accentuated by good quality, delight, good experience, and extra value-added features. In order for the companies to fulfill these characteristics, it is hereby suggested to use "mind mining" methods

to identify customers' unsatisfied and latent needs. They also ought to use methods and means to eliminate the inconvenient utilization, and to integrate product functions and quality. These findings and results may be hereby included into the quality functions. For realizing these quality functions, the drivers are the capability of "technology innovation." As such, the four constructs form a "transfer loop of the innovation system," as shown in **Figure 5**.

8. Conclusion

In a highly globalized economy, it is widely recognized that "innovation" has become the key force for achieving business excellence, which in turn reflects as competitive advantages, growth, and development [66]. Only the pursuit of product quality and service quality is not enough to achieve the business objectives. The enterprises should alter their quality concepts from the narrow definition of quality to include customer value and innovation. In order to realize the innovation performance, this study has proposed a framework of the innovation system and the related transfer loop of the system. However, the proposed system is a conceptual model, empirical studies ought to be conducted to identify potential causalities among the critical constructs and their practices.

Author details

Ching-Chow Yang

Address all correspondence to: chinchow@cycu.edu.tw

Department of Industrial and Systems Engineering, Chung Yuan Christian University, Taiwan, Republic of China

References

[1] Ulwick, Anthony W. *What Customers Want.* 2005; New York: McGraw-Hill Inc.

[2] Yang, C.-C. The Integration of TQM and Six-Sigma. In the book "*Total Quality Management and Six Sigma*", edited by Tauseef Aized, ISBN 978-953-51-0688-3, InTech Ltd, August 8, 2012; 219–246.

[3] Edvard, C. D. The meaning of quality. *Quality Progress*. 1968; 1(10), 36–39.

[4] Deming, W. E. *Out of Crisis.* 1986; Cambridge, MA: MIT Press.

[5] Gorst, J., Kanji, G., & Wallage, W. Providing customer satisfaction. *Total Quality Management*, 1998; 9(4–5), 100–103.

[6] Sirohi, N., McLaughlin, E.W., & Wittink, D. R. A model of customer perception and store loyalty intentions for a supermarket retailer. *Journal of Retailing*, 1998; 74(2), 223–245.

[7] Teece, D. J. Business model, business strategy and innovation. *Long Range Planning*, 2010; 43, 172–194.

[8] Rindova, V. P., & Petkova, A. P. When is a new thing a good thing? Technological change, product form design, and perceptions of value for product innovations. *Organization Science*, 2007; 18(2), 217–232.

[9] Reichheld, F. F. Loyalty and renaissance of marketing. *Marketing Management*, 1994; 2(4), 10–21.

[10] Schneider, B., & Bowen, D. E. Understanding customer delight and outrage. *Sloan Management Review*, 1999; 41(1), 35–45.

[11] Kumar, A., Olshavsky, R. W., & King, M. F. Exploring alternative antecedents of customer delight. *Journal of Customer Satisfaction, Dissatisfaction and Complaining Behaviour*, 2001; 14, 14–26.

[12] Gallo, C. *The Apple Experience.* 2012; New York: McGraw-Hill Inc.

[13] Yang, C.-C. The identification of customer delight for quality attributes and its applications. *Total Quality Management & Business Excellence*, 2011; 22(1–2), 83–98.

[14] Rao, A., Carr, L. P., Dambolena, I., Kopp, R. J., Martin, J., Rafii, F., & Schlesinger, P. F. *Total Quality Management: A Cross Functional Perspective.* 1996; New York: John Wiley & Sons.

[15] Zokaei, K., & Simons, D. Performance improvements through implementation of lean practices: A study of the U.K. Red Meat Industry. *International Food and Agribusiness Management Review*, 2006; 9(2), 30–53.

[16] Yang, C.-C. Improving the definition and quantification of quality costs. *Total Quality Management & Business Excellence*, 2008; 19(3–4), 175–191.

[17] Mitra A. *Fundamentals of Quality Control and Improvement Second edition.* 1998; London: Prentice-Hall International (UK) Limited.

[18] Xie, M., Goh, T. N., & Ranjan, P. Some effective control chart procedures for reliability monitoring. *Reliability Engineering & System Safety*, 2002; 77(2), 143–150.

[19] Lavin, M. Inspection efficiency and sampling inspection plans. *Journal of the American Statistical Association*, 1946; 41(236), 432–438.

[20] Giakatis, G., Enkawa, T. & Washitani, K. Hidden quality costs and the distinction between quality cost and quality loss, *Total Quality Management,* 2001; 12(2), 179–190.

[21] Oakland, J. S. *Total Quality Management.* 2nd ed. 2000; Jordan Hill, Oxford: Butterworth-Heinemann.

[22] Juran, J. M. (1974).*Quality Control Handbook.* 3rd ed. New York: McGraw-Hill.

[23] Feigenbaum, A. V. Total Quality Control. *Harvard Business Review*, 1956; 34 (6), 93–101.

http://dx.doi.org/10.5772/67211

[24] Crosby, P. *Quality is Free*. 1979; New York: McGraw-Hill.

[25] Dotchin, J. A., & Oakland, J. S. Theories and concepts in total quality management. *Total Quality Management*, 1992; 3(2), 133–145.

[26] Powell, T. C. Total quality management as competitive advantage: A review and empirical study. *Strategic Management Journal*, 1995; 16, 15–37.

[27] Elshennawy, A. K., Maytubby, V. J., & Aly, N. A. Concepts and attributes of total quality management. *Total Quality Management*, 1991; 2(1), 75–97.

[28] Yang, C.-C. An integrated model of TQM and Six-Sigma. *International Journal of Six Sigma and Competitive Advantage*, 2004; 1(1), 97–111.

[29] Sila, I., & Ebrahimpour, M. An investigation of the total quality management survey based on research published between 1989 and 2000. *Journal of Quality and Reliability International*, 2002; 19(6–7), 902–970.

[30] Pande, P. S., Neuman R. P., & Cavanach R. R. *The Six Sigma Way*. 2000; New York: McGraw-Hill.

[31] Antony, J., & Banuelas, R. Key ingredients for the effective implementation of Six Sigma program, *Measuring Business Excellence*, 2002; 6(4), 20–27.

[32] Bell, R. & Keys, B. A conversation with Curt W. Reimann on the background and future of the Baldrige award. *Organizational Dynamics*, 1998; 26(4), 51–61.

[33] Chin, K. S., Pun, K. F., & Hua, H. M. Consolidation of China's quality transformation efforts: A review. *International Journal of Quality and Reliability Management*, 2001; 18(8), 836–853.

[34] Hendricks, K. B., & Singhal, V. R. Quality awards and the market value of the firms: an empirical investigation. *Management Science*, 1996; 42(3), 415–436.

[35] Gunasekaran, A. Enablers of total quality management implementation on manufacturing: A case study. *Total Quality Management*, 1999; 10(7), 987–996.

[36] Hansson, J., & Eriksson, H. The Impact of TQM on financial performance, *Measuring Business Excellence*, 2002; 6(4), 44–54.

[37] Youssef, M. A., Boyd, J., & Williams, E. The impact of total quality management on firms responsiveness: An empirical analysis. *Total Quality Management*, 1996; 7(1), 127–144.

[38] EFQM. *The European Model for Total Quality Management*, 1992; Brussels: European Foundation for Quality Management.

[39] Tan, K. C., Wong, M. F., Mehta, T., & Khoo, H. H. Factors affecting the development of national quality awards. *Measuring Business Excellence*, 2003; 7(3), 37–45.

[40] Cauchick Miguel, P. A., Morini, C., & Pires, S. R. I. An application case of the Brazilian National Quality Award. *The TQM Magazine*, 2004; 16(3), 186–193.

[41] Yang, C.-C. Development of an integrated model of a business excellence system. *Total Quality Management & Business Excellence*, 2009; 20(9), 931–944.

[42] Kanji, G. K. Architecture of business excellence in the public and service sectors. *Total Quality Management and Business Excellence*, 2008; 19(3–4), 399–415.

[43] Okumura, S., Morikuni, T., & Okino, N. Environmental effects of physical life span of a reusable unit following and physical failures in a remanufacturing system. *International Journal of Production Research,* 2003; 43(16), 3667–3687.

[44] Zairi, M. Beyond TQM implementation: the new paradigm of TQM sustainability. *Total Quality Management & Business Excellence*, 2002; 13(8), 1161–1172.

[45] Yang, C.-C. Exploration strategies and key activities for the system of environmental management. *Total Quality Management & Business Excellence,* 2011; 22(11), 1179–1194.

[46] Matzler, K., & Hinterhuber, H. H. How to make product development projects more successful by integrating Kano's model of customer satisfaction into quality function development. *Technovation*, 1998; 18(1), 25–38.

[47] Finkelman, D. P., & Goland, A. R. How not to satisfy your customers. *The McKinsey Quarterly*, 1990; 26(4), 2–12.

[48] Heskett, J. L., Jones, T.O., Loveman, G.W., Sasser, W. E., & Schlesinger, L.A. Putting the service-profit chain to work. *Harvard Business Review*, 1994; 72(2), 164–174.

[49] Jones, T. O., & Sasser Jr., W. E. Why satisfied customers defect. *Harvard Business Review*, 1995; 73(6), 88–99.

[50] Johnson, R., Retail isn't broken Stores are. *Harvard Business Review*, 2011; 89(12), 78–82.

[51] Johann, F., & Kurt, M. Customer delight and market segmentation: An application of three-factor theory of customer satisfaction on lift style groups. *Tourism Management*, 2008; 29(1), 116–126.

[52] Torres, E. N., & Shery, k. From customer satisfaction to customer delight: Creating a new standard of service for the hotel industry. *International Journal of Contemporary Hospitality Management*, 2013; 25(5), 642–659.

[53] Rigby, D. K., Gruver, K., & Allen, J. Innovation in turbulent times, *Harvard Business Review*, 2009; 87(6), 79–86.

[54] Moore, G. A., Innovating within established enterprises. *Harvard Business Review*, 2004; 82(7–8), 87–92.

[55] Yang, C.-C., The establishment of a TQM system for the health care industry. *The TQM Magazine*, 2003; 15(2), 93–98.

[56] Hammer, M., & Champy, J. *Reengineering the Corporation.* 1993; New York: Harper Collins.

[57] Hendricks, K. B., Singhai, V. R. & Stratman, J. K. The impact of enterprise systems on corporate performance: A study of ERP, SCM, and CRM system implementations. *Journal of Operations Management*, 2007; 25(1), 65–82.

http://dx.doi.org/10.5772/67211

[58] Bill, G. *Business @ the Speed of Thought: Using a Digital Nervous System.* 1999; New York: Warner Books, Inc.

[59] Chuck, M. *Get Future*. 1999; New York: McGraw-Hill, Inc.

[60] Feigenbaum, A. V. The new quality for the twenty-first century. *The TQM Magazine*, 1999; 11(6), 376–383.

[61] Ulwick, A. W. Turn customer input into innovation. *Harvard Business Review*, 2002; 80(1), 91–97.

[62] Yang, C.-C. An analytical methodology for identifying the latent needs of customers. *Total Quality Management & Business Excellence*, 2013; 24(11–12), 1332–1346.

[63] Thomke, S., & von Hippel, E. Customers as Innovators: A new way to create value. *Harvard Business Review,* 2002; 80(4), 74–81.

[64] Hurley, R. F., & Hult, G. T. M. Innovation, market orientation and organizational learning: an integration and empirical examination. *Journal of Marketing*, 1998; 62(3), 42–54.

[65] Hult, G. T. M., Hurley, R. F., & Knight, G. A., Innovativeness: Its antecedents and impact on business performance. *Industrial Marketing Management,* 2004; 33, 429–438.

[66] Sohel-Uz-Zaman, A. S. M. D., & Anjalin, U. Evolution of service: Importance, competitiveness and sustainability in the new circumstances. *Journal of Service Science and Management*, 2011; 4, 253–260.

Chapter 2

Key Aspects for Implementing ISO/IEC 17025 Quality Management Systems at Materials Science Laboratories

Rodrigo S. Neves, Daniel P. Da Silva,
Carlos E. C. Galhardo, Erlon H. M. Ferreira,
Rafael M. Trommer and Jailton C. Damasceno

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66100

Abstract

Implementing a quality management system based on the requirements specified in ISO/IEC 17025 standard at materials science laboratories is challenging, mainly due to two main factors: (i) the high technical complexity degree of some tests used for materials characterization and (ii) the fact that most materials science laboratories provide materials characterization tests and also carry out research and development activities. In this context, this chapter presents key subjects while implementing a quality management system at materials science laboratories and some considerations on strategies for effectively implementing such systems.

Keywords: quality management, quality assurance, materials metrology, ISO/IEC 17025, materials science laboratories

1. Introduction

The constant strive for innovation on the development of new products and services led, in the last decades, to a growing interaction between research and development (R&D) laboratories. Most of these laboratories are from universities and research institutes and interact with diverse branches of productive sector. As a consequence, several of them opted for implementing a quality management system based on an international requirement standard aiming at improving products and services reliability [1–6].

Particularly interesting cases are the materials science laboratories [7]. Many of these laboratories, originally headed almost exclusively to R&D activities, became interested in providing accredited tests for materials characterization. As a consequence, they should adopt a quality management system based on the ISO/IEC 17025 requirements standard [8]. On the other hand, materials characterization laboratories specialized in providing accredited high-technology tests are constantly pushed to carry out R&D, in order to improve or develop new processes and services. One way or another, both cases describe scenarios where quality management system must run in an environment where R&D and accredited testing services co-exist.

Considering this context, implementing quality management systems at materials science laboratories is a challenging task due to several reasons. The first point is the fact that whilst the implementation of an ISO/IEC 17025 oriented quality management system at testing laboratories is a consensus, the advantages and disadvantages of applying quality management concepts to R&D activities are still issues widely discussed by the quality management community. Indeed, the main aspects of this controversy are highlighted in the articles of Mathur-De-Vré [9] and Krapp [10]. They state that adding quality management concepts, like metrological traceability, trackability and rational human resources management, must not restrain the flexibility necessary in order to carry out R&D activities.

In spite of this contradiction, materials science laboratories, but also R&D laboratories of several scientific areas, start to see the standardization promoted by the implementation of a quality management system based on ISO/IEC 17025 standard as a way to reduce costs, reduce tests execution deadlines, satisfy customers and especially add quality to their services and products. More recently, an article published in *Nature* has drawn attention to how quality management and quality assurance concepts could improve R&D activities and results [11].

In this context, this chapter aims to discuss practical aspects for implementing a functional quality management system at materials science laboratories. Considering this approach, in the reminder of this section a brief discussion on the basic principles of ISO/IEC 17025 will be presented, followed by a practical approach related to the main aspects of quality management system implementation.

The aspects discussed in the aforementioned method are mainly associated with the experience of the authors regarding the implementation of an ISO/IEC 17025 oriented quality management system at the Materials Metrology Division (Dimat) of the National Institute of Metrology, Quality and Technology (Inmetro), the Brazilian National Metrology Institute (NMI).

1.1. ISO/IEC 17025—General requirements for the competence of testing and calibration laboratories standard: basic principles

ISO/IEC 17025 specifies the basic requirements for the competence verification of laboratories carrying out testing and calibration activities, focused in meeting customer expectations and keeping organized laboratory records and documents. These requirements relate to most, if not to all, the laboratory activities concerning testing and calibration services provided by the laboratory, from the control of documents and records to the technical procedures standardization. The standard separates these requirements into two wide classes: management

http://dx.doi.org/10.5772/66100

requirements and technical requirements. The former will be briefly described in this section, since complete information can be found in the ISO/IEC 17025.

Management requirements are mainly related to organization, control and update of documents, analysis of contracts and monitoring optimization of the quality management system. Its main topics are summarized as follows:

- Management: this item includes primarily a definition of the company/laboratory organization, describing the quality management and technical team, the attribution of responsibilities and the clear commitment to the quality management principles;
- Control of documents and records: this section incorporates the design of a system to control the revision of documents and ensure access to up-to-date documents, avoiding use of obsolete versions. It also encompasses a system to records control;
- Review of requests and contracts: review of tenders, requests and contracts must cover all the contracts with clients and with suppliers and subcontracted service providers, in as much as these are related to the activities covered by the quality management system;
- Non-conforming control: it includes identification of the non-conforming cause, its corrections, the application of further corrective actions when necessary and prevention of potential non-conformities. It also includes long-term monitoring of corrective actions effectiveness;
- Contact with the client: this point counts all the contacts with clients. It comprises on clients' complaints;
- Monitoring and optimization of quality management system: this item covers several actions, like a programme that monitors and collects information on the effectiveness of the quality management system. It equally encapsulates critical analysis of such information in order enhance the system's optimization.

On the other hand, technical requirements are directly related to testing and calibration procedures, as summarized below:

- Technical staff: covers mainly the evidence of technical staff qualification with respect to the assigned tasks;
- Control of environmental conditions and accommodations: covers adequacy of accommodations, including its organization, to activities carried out at the laboratories. Furthermore, if testing or calibration activities demand specific environmental conditions (temperature and humidity) such conditions must be controlled;
- Test and calibration methods: comprises all the aspects related to the methods used in the laboratory, like evidence of the use of standard methods and validation of non-standard and laboratory-developed methods. It also includes the evaluation of uncertainty of measurements;
- Equipment: contains identification and guaranties of working conditions and, when necessary, calibration of the laboratory equipment;

- Measurement traceability: deals with traceability via calibration or use of certified reference materials or standards;
- Sampling: involves the use of standard sampling methods adequate to different sort of materials;
- Handling of testing and calibration items: covers definition of procedures to identify and handle samples to tests, as well as records of trackability;
- Results quality assurance: covers a programme of actions related to the quality assurance of obtained results. Such a programme may contain actions, for example, participation in inter-laboratory comparison or proficiency-testing programmes; and
- Reporting results: this topic consists of the procedures required for communicating tests and calibration results to the clients.

Up to this date, ISO/IEC 17025 standard was last revised and confirmed by ISO in 2010 and is currently under a new revision process [12]. Accordingly to some stakeholders, several aspects should be reviewed in order to modernize the standard concepts and specifications, as, among others [13] quality system management designing based on performance and process, trackability, modernization of requirements related to software and electronic records and the evaluation of other standards (as ISO/IEC 17065—conformity assessment—requirements for bodies certifying products, processes and services, for example) relevance to ISO/IEC 17025.

Although these requirements are separated into two main classes, they should not be seen as isolated. As a matter of fact, it is easier to see the relation between them if one considers the basic system development concept underneath the implementation of a quality management system based on a requirement standard. This concept is represented in **Figure 1**.

Any quality management system such as the ISO/IEC 17025 is based on a tripod formed by policies, procedures and records.

Policies are the platform for the development of the system. They must clearly dispose the figures responsible for the different processes that constitute the management system itself, as well as the basic principles that encompass these processes. As an example, the policy for documents control may dispose that update and access control is the responsibility of a centralized system (a document control centre) for this very purpose, whilst the production of technical documents is the responsibility of the laboratory staff. Based on this simple policy, documents control centre and laboratories can, together, develop procedures to ensure the correct access, update and utilization of documents.

As per ISO-specific nomenclature, all policies must be documented. On the other hand, procedures describe how a quality management process must be done and how to register its execution. They apply to most of the processes carried out in a laboratory, stretching from technical procedures for testing and calibration to procedures about documents and records production, contracts and suppliers' evaluation, and quality assurance procedures, to name but a few.

http://dx.doi.org/10.5772/66100

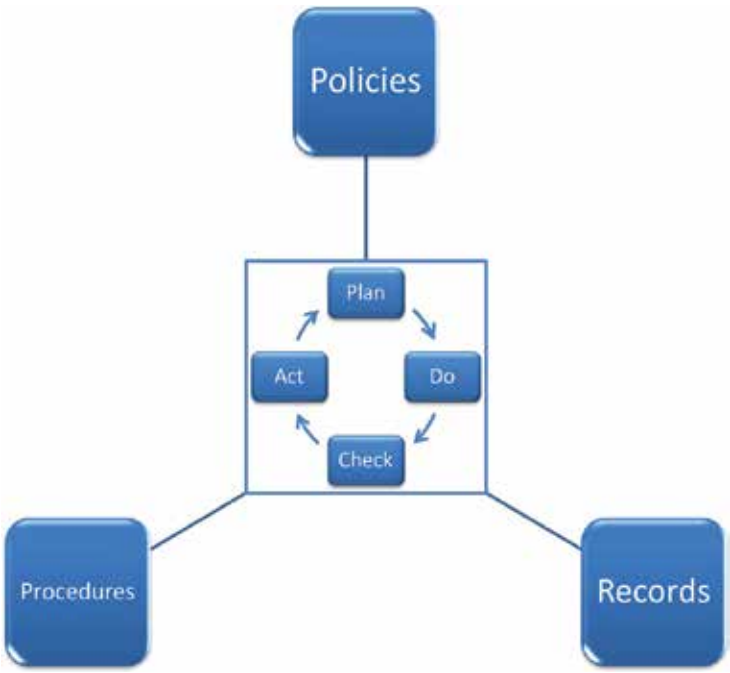

Figure 1. Basic concepts underneath a quality management system based on a requirements standard.

Finally, it is worth to mention that the records comprise not only testing and calibration results, but also complete status of laboratory equipment, staff training and ambient conditions. In other words, the records consist of all activities related to the quality management system. The way the recording is performed, stored and made available is usually specified in procedures. Thus, all requirements should be considered in unison so as to design a functional quality management system.

One last, but fundamental aspect related to the policies-procedures-records tripod is that although it constitutes the base (policies) and the tools (procedures and records) for the quality management system operation, the system itself is under constant improvement. In this way, procedures and even policies are constantly evolving in order to simply correct eventual non-conforming or to optimize the system. In **Figure 1**, this characteristic is depicted by the central cycle representing the 'plan-do-act-check' system (PDCA).

All the requirements disposed in ISO/IEC 17025 must be fulfilled when implementing the quality management system. However, the extensive discussion of each one of the requirements is not the intent of this chapter. Instead, this chapter will cover the aspects regarded as key requirements for implementing a functional and flexible quality management system at materials science laboratories. Section 3 addresses some practical topics concerning strategies for implementing a quality management system and, lastly, final considerations are presented in Section 4.

2. Key matters in quality management systems in materials science laboratories

2.1. Organization of laboratories: the quality management system set out

Implementation of a quality management system at materials science laboratories involves a clearly specified chronogram. Additionally, responsibilities of the laboratory staff members should be defined in detail in order to organize the process.

However, a critical question concerning the necessary units for work organisation whilst implementing a functional and lissom quality management system must be answered.

Most of materials science laboratories make use of several experimental techniques to characterize different materials properties, such as structure, chemical composition, and thermophysical and mechanical characteristics among others. Therefore, as the scope of techniques associated with materials science covers a wide area, it is not rare that the same organization (company or laboratory) holds multiple analysis techniques that may differ from applications, scale or sort of materials. In addition, each one of them demands very specific requirements concerning facilities and staff.

One possible approach is to gather similar experimental techniques in the specific laboratory units. The quality management team, in co-operation with the technical staff, may choose similar and/or complementary techniques to compose each unit. Another approach is to organize the laboratories driven by external demand. For example, an institute/company for testing building materials may organize multiple laboratories following the regulatory compliance, such as energy efficiency, thermal comfort and fire safety. Another common case is that of laboratories inside companies producing materials, designed for materials quality control. In such cases, the laboratories can be organized according to specific production needs.

On the other hand, not only external demands, but also internal factors should be considered. The staff size must be compatible with the experimental complexities and tests demands. Besides, physical and financial aspects are also important, because a laboratory divided into multiple locations without financial autonomy may result in an unnecessarily complicated management system that could hamper the management.

Furthermore, if the laboratory belongs to a company or may form another legal entity, the organizational structure must ensure that duties are clearly assigned, eliminating the possibility of conflict of interests between the laboratory and any other operating units.

Considering these aspects, it is the authors' view that a rational organization of technical infrastructure and equipment in a minimal, but necessary, number of different laboratories or units can minimize financial and human resources costs. If the structure is spread among many different laboratories, the production of documents may turn into a burden, resulting in a highly complex and oversized quality management system, presenting, for instance, redundant procedural documents in different laboratories. Otherwise, the distribution of the structure between few laboratories may produce a frame where many different techniques are condensed in a single unit, creating very complex laboratories which will be hard to manage.

http://dx.doi.org/10.5772/66100

Besides the practical aspects presented above, it is equally important to understand how the customer sees the company's activities. As an example, consider a client who wants tests for compliance to some specific standard or regulation, as is the case of characterization of thermal conductivity for insulator materials. In this case, it will be much easier to maintain a quality system and to organize records and documents when laboratories are directly related to the sets of tests performed for each regulation. However, if the customers see the organization as a centre of excellence in materials science, organizing laboratories by similar techniques could be more appropriate.

As one can see, the approach adopted to organize the technical infrastructure in laboratories or units is a 'tailor made' task that impacts several aspects of the strategy used for implementing the quality management system. It determines the background for the definition of policies and processes common to all laboratories, such as documents control, services and supplies acquisition, customer service, internal audit, control of records and management of nonconforming works and events.

After that technical requirements specific to each laboratory/unit could be easily implemented by combining particularities of each one of them with the common policies and processes, in order to optimize system standardization. In this aspect, a well-designed organization of laboratories/units automatically shows that keeping some processes common to all laboratories minimizes the efforts from technical staff in document production and also improves the collaborative process for quality management.

2.2. Control of documents and records

The purpose of a document control system is to deal with the large amount of documents, such as procedures, reports and forms generated by the quality management system. In such a system, any given document must be considered as a unique element. It must have a unique identifier, like a name or code, and each one of these documents has a lifetime in the system, from its release to its cancel dates. The documents *per se* evolve over time, necessitating thus, a revision. Therefore, a functional documents control system must keep track documenting the corresponding characteristics.

The system must control not only the internal procedures developed within the company or laboratory, but it must also include instrument manuals, software, drawings, standards and regulations as well as external source documents (i.e. national or international standards for testing of materials). This last class of documents usually demands special attention, since the revision of such standards by the responsible organizations must be periodically monitored by the laboratory, in order to avoid the use of obsolete external standards.

In practice, the complexity of such a system depends on each case. For small laboratories carrying out a small number of tests, an electronic spread sheet could be enough for document management. However, for larger corporations with several laboratories, a system running a consolidated documents database allowing relational searches could be a proper choice. Apart from its complexity, the system must keep track of a certain number of document characteristics such as:

- release date;
- revision date (from the first to the later revision);
- unique identification for each revision (revision number); and
- copies of all revisions, authors, reviewers, and the person(s) who approved each document.

Additionally, invalid and obsolete versions of documents must be clearly identified in the system, in order to avoid misuse. Finally, the system must provide access to all its documentation to the users.

Document management goes beyond an effective document control system. The standardization and documentation of processes is a tool to make activities much more effective and efficient. Documents should be easy to read and simple to use. They must disseminate knowledge within the organization. They could be hierarchical and should be referred to each other in such a way that it improves the document writing process by reducing the amount of rework. They should be based on consensus of those involved in the process. Each document must describe the process accurately, making use of figures, tables and flowcharts as and where necessary. They must be written with its practical use in mind and focused on being easily understood.

Records must be also included in a management system. Several records must be kept at the laboratory level—such as testing results—and must be organized considering the particularities of each laboratory or unit.

There are other records that can be kept organized outside the laboratory to provide easy access, as follows:

- meeting minutes,
- management review minutes,
- implementing reports,
- item registration,
- personnel documentation and
- test item handling records.

The record management system may also keep track of measures of productivity, assisting management decisions. It is the authors' view that control of records may impose bottlenecks when a quality management system is applied to R&D activities. Indeed, during R&D projects execution, detailed investigation of a considerable number of parameters is carried out. These activities result in a large number of preliminary results that will not be included in the research outcome. However, even these preliminary results must be accounted as controlled records by the quality management system.

All these aspects make designing a system to control documents and records a customized task, which must take into account specific characteristics of the company/laboratory, as the

http://dx.doi.org/10.5772/66100

mechanism to provide rapid access to documents and eventual access privileges. However, once a system of documents and records management is working properly, it makes laboratory-daily activities and follow-up tasks easier, since the information is always organized and readily available.

2.3. Customer relationship service

Customer service is a common requirement for all testing laboratories and the usual approach perfectly applies to materials science laboratories. When a company or institution structure consists of several laboratories, the usual approach can be applied. Establishing a specialized sector for customer service is a suitable policy that allows laboratories technical staff to spend more time for laboratories activities (testing services and R&D activities). Additionally, by minimizing the direct contact between the client and the technical staff, the customer service sector may also avoid or at least reduce some unpleasant situations, like external pressure for anticipating deadlines or questions about price. The customer service sector should be assigned to handle all the communication between the company and the clients, from the first commercial contact to the finalization of the contracted service. It is not uncommon that the same channel used for customer service also fulfils the function of customer's claims channel. Although these observations are pretty obvious, it is addressed here because it is crucial for R&D laboratories wishing to start providing accredited tests to properly take care of costumer services from the start.

Although customer service sector limits the direct contact between the client and the technical staff, in the case of materials science laboratories this direct contact should be done in order to clarify technical aspects related to the test as demanded by the costumer. These technical aspects normally relate to the use of specific experimental procedures and an eventual deviation from standard testing procedures in order to suit customer needs. It may also be due to some specific characteristic of the tested material. It is important to mention that when the laboratory uses methods other than standardized ones, even due to customer's request, such methods must be validated by the laboratory technical staff.

2.4. Calibration, traceability to the SI and quality assurance of test results

The most relevant impact of implementing a quality management system in any laboratory developing tests and R&D is the quality assurance of the experimental results. This assurance is important not only to a client interested in materials properties characterization, but also to the R&D activities, since traceability and reproducibility of results are deemed as key factors in science and technology. In this aspect, the implementation of a quality management system based on the ISO/IEC 17205 shall demonstrate confidence in experimental results. From a technical standpoint, the establishment of calibration and traceability to appropriate measurement standards is fundamental, and consequently the quality assurance can be demonstrated. As such, this section discusses how calibration and traceability can be handled in materials science laboratories in order to improve quality assurance of experimental results.

ISO/IEC 17025 standard requires that all equipment used for testing or calibration must be calibrated in order to assure the quality of results. This applies also to subsidiary measurements, such as environmental conditions monitoring, as long as they have a significant contribution to the total uncertainty of the results. An useful example is the International Vocabulary of Metrology (VIM) [14], at page 28, defines calibration as '… operation that, under specified conditions, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication'. The purpose of the calibration is not simply to ensure that the equipment is providing a correct result, but most of all, to guarantee a link of this measurement with the unit definitions in the international system of units (SI). This link, that establishes a relation between the result of a measurement and the abstract SI units, is called metrological traceability.

This traceability, however, can also refer to a measurement procedure including the measurement unit or a measurement standard. The concept of traceability is implied when dealing with traditional calibration procedures, such as the calibration of a scale. In this case, a standard weight is used to calibrate the scale in a simple procedure. Therefore, traceability is secured by an unbroken chain of calibrations tracing back to the international mass standard, which is unique and is the physical realization of the kilogram unit. This simple procedure, however, is far from the reality of an engineering materials testing laboratory.

When dealing with tests of materials, some of their properties may not be directly linked to the SI through a simple metrological traceability chain [15]. Typically, even a simple measurement of a material property involves many different measurements and links to several SI units to guarantee the metrological traceability. An useful example is the thermal conductivity, whose unit is [W/m K] or equivalently [m kg s^3 K], measured by the GHP method. In this case, the metrological traceability to four different units needs to be provided, which in turn, requires calibration of several instruments used in the process. Although this is the primary method to establish the traceability to SI, this is a very laborious task and, in some cases, it is hard to be used for some materials characterization techniques. In these cases, alternative approaches are used [16].

Another topic that makes testing of materials not so easy task is that some properties of a material are not intrinsic but dependent on the measurement procedure, which is called a procedural property. This means that the measurement of a given material property can return different values depending on the measurement method. A typical example is the hardness. In this case, the traceability is guaranteed by not only the calibration of the instruments but also by a documented standard defining the method. Finally, sometimes a measurement cannot be linked to SI, as a result of the measurement is not an SI unit but a classification on an appropriate numerical scale. Once again, the metrological traceability in this case is given by a measurement standard procedure and some reference materials. An example is the Mohs Hardness scale.

http://dx.doi.org/10.5772/66100

In all cases, the ISO/IEC 17025 standard gives alternatives to foster confidence in measurements by establishing the traceability to appropriate measurement standards. Those alternatives are as follows:

- the use of certified reference materials (CRM) provided by a competent supplier;
- the use of specific methods and/or consensus standards; and
- participation in suitable inter-laboratory comparisons where possible.

For materials testing, the use of certified reference materials (CRM) is typically the preferred choice. It is important that the reference material should be accompanied by a document, issued by a recognized organization. The latter provides values of specific properties with associated uncertainty and metrological traceability, using validated procedures. It is worth mentioning at this point that, when dealing with a specific material, there are many of its properties that can be assessed and used as a reference in a measurement, but the CRM is typically certified for just one of its properties of interest.

The certification of this property can be achieved through a measurement using a primary method, with the lowest uncertainty and with the necessary traceability. This reference material is then used to provide the traceability to the measurements of this property in a much simpler way. The drawback is that the measurement uncertainty is a bit higher than the one obtained by primary methods, but most of the time this is low enough for the researcher's purpose.

There are still cases where one cannot certify a reference material using a primary method. In this case, the certification is achieved by consensus. This means that different laboratories make the measurement using a given protocol, and the mean value is used as a reference value for that given property.

A laboratory participating in such a comparison can (or cannot) demonstrate their competence and measurement capability, without having to provide traceability to SI for their results. This is similar to laboratories that realize the SI units, and therefore, have no physical standard that can be used to calibrate their instrument. Those laboratories undergo international inter-laboratory comparisons with other laboratories in the same status, in order to mutually guarantee their measurement capabilities. In the specific case of national metrology institutes, there is the possibility of participation in key comparisons organized by the Bureau International de Poids et Mesures (BIPM) [17].

If there is no CRM available or, if no inter-laboratory comparisons have been performed, an alternative way to assure the quality of test and calibration is to periodically perform intra-laboratory comparisons. In an intra-laboratory test, two or more researchers perform the material measurement, using—preferentially—a CRM and calibrated equipment, and the results are compared. With this practice, it is possible to analyse data and detect trends that can influence the result and uncertainty.

At BIPM there are several consultative committees (CC) that are responsible for the base SI units, namely:

- length,
- mass,
- time,
- electric current,
- thermodynamic temperature,
- amount of substance and luminous intensity.

Indeed, the same applies for the derived units or for specific fields such as acoustics, ultrasound and ionizing radiation. There is, so far, no specific CC for materials. BIPM understands that any material property can be treated under one of the current CCs. For example, thermal conductivity is treated under the consultative committee for thermometry (CCT).

An important forum for discussion on measurement methods and procedures are the ISO technical committees (TC). Once again there is no specific TC for materials metrology, but there are so many TCs; it is unlikely that one given property of a material is not treated in one of the committees. ISO is a normative organization and as such it is concerned with publishing standards, either for a material testing or for method standardization. Typically, a normative standard is produced when there is already a well-established consensus in the validity of such a method or test that one wants to standardize. However, one can propose inside a TC a study group to develop and validate a new method before writing the standard, taking advantage of the internationality and wide extent of the members.

This procedure is however not preferable, since it may take too long to fully develop new methods, and the ISO time is short and well defined.

A solution to this problem is the pre-normative forums. One of the most important forums is the Versailles project on advanced materials and standards (VAMAS), whose main objective is to '…promote world trade by innovation and adoption of advanced materials through international collaborations that provide the technical basis for harmonization of measurement methods, leading to best practices and standards' [18]. The VAMAS has been founded under the auspices of the G7 economic summit in 1982 and includes a number of representatives from many countries in the world. The VAMAS consists of the technical working areas (TWA) dealing with a class of materials or a given property of materials, such as polymer composites, mechanical properties of thin films and nanoparticle populations. In every TWA, there are many projects that deal with specific subjects. In this way, one can propose new measurement methods or new testing for new materials, run inter-laboratory comparisons with the collaboration of different national metrology institutes and also other stakeholders from industry and academia.

The VAMAS is also liaising with the BIPM in the development of materials metrology. While VAMAS can propose new procedures and measurement methods, BIPM would run the inter-laboratory comparisons among the NMIs using their well-established system of key comparisons. Finally, if the project is successful, one might forward it to the ISO and develop a standard out of it, since it has already been tested by a worldwide community.

http://dx.doi.org/10.5772/66100

In order to manage equipment and reference materials, calibration routine and records, it is critical to have a calibration programme. In such a programme, the laboratory shall control the calibration status of the equipment and the reference materials that are relevant for the testing procedure. Different equipment presents different needs regarding its calibration periodicity. For instance, a scale or ruler can be calibrated on an annual basis, because their daily use does not change their calibration condition, if well conserved and appropriately handled. On the other side, some types of apparatus may require a calibration before the beginning of each measurement, as their calibration status may not be valid for a long period. Finally, for multi-faceted equipment used in typical material testing laboratories, their calibration is too complex to be done on a periodical basis. One must, however, confirm that performance properties or legal requirements of the measuring system are achieved. This is called verification. When possible, both calibration and verification of the equipment should be done with the use of reference standards or certified reference materials. It is also quite important to verify the status of any apparatus when it goes through maintenance or when it is used outside the laboratory's environment.

Not only equipment, but reference standards and CRM must also have a systematic control of their calibration situation or validity of its certification. While reference standards can be re-calibrated on a periodic basis just like any equipment, CRMs normally have a short life span that guarantees their properties. After its expiration, the laboratory must then acquire new CRMs or, when possible, get its re-certification. Particularly to some materials, due to their stable properties, the validity of their certification can be indefinite. As an example hereto, the alumina powder (corundum) reference material for quantitative phase analysis using the powder diffraction method is mentioned. This reference material has been developed by the NIST [19] whose certification is valid indefinitely as long as it is stored and handled according to its certificate.

Additionally, in some particular cases, it is not possible to calibrate the instrument and there is no CRM available for instrument verification. In such cases, ISO/IEC 17025 recommends to repeat the tests using samples already tested and retained in the laboratory and/or to compare the results of tests performed using different methods, in order to provide some evidence of quality assurance.

The routinely monitoring of calibrations or verifications can be recorded with the use of control charts that provide an easy way to examine the overall status of the equipment or reference standards and identify eventual trends that may affect the results of tests. In order to assure the quality of results, it is important that the laboratory defines and documents the acceptance criteria of the calibrations of equipment, reference standards, and the certificates of CRMs. The control charts are viewed as a very valuable tool for this task. It is worth mentioning at this point that activities related to quality assurance not only involve technical aspects, but also organizational elements that permeate work at all stages. Thus, a quality system with good documentation, internal and external auditing and records (equipment maintenance, corrective/preventive actions, etc.) strongly supports quality assurance.

2.5. Evaluation of uncertainty of measurements

Measurement uncertainty is a quantitative indication of quality for measurement results. When the uncertainty related to a result is not declared, this result cannot be compared with specified reference values or standards. As a matter of fact, even results from different tests related to the same material/sample are hard to compare without evaluation of uncertainty. Uncertainty evaluation is essential to guarantee the metrological traceability of measurement results and to ensure they are accurate and reliable. Additionally, measurement uncertainty must be accounted for whenever a decision has to be taken based on measurement results.

In this context, progressive globalization of markets pushed for the use of a standard procedure for evaluating uncertainty of measurements, in order to assure comparability of results and, consequently, mutual recognition in metrology. In this aspect, laboratories accredited under the ISO/IEC 17025 standard aiming to demonstrate their technical competence and the ability to properly operate their management systems are required to evaluate uncertainty of their measurement results. The two main approaches commonly used for evaluating measurement uncertainty are the bottom-up and the top-down and both of them will be briefly discussed here.

The bottom-up approach is more often used for classic physical metrology systems, such as mass or dimension measurements. It involves using a model equation (the one used for the calculation of the measurand) as a starting point and then considers all individual uncertainty contributions. This is the approach that is well described in the guide to the expression of uncertainty in measurement (GUM) [20] and it uses the law of propagation of uncertainties as its base. The GUM is a guide to uncertainty evaluation and it is the more widely accepted approach for uncertainty evaluation in metrology studies regarding the comparison of results.

Considering the approach described in GUM, the measurand is described as a function of several variables called input quantities. For each input quantity, the sources of uncertainty should be evaluated, quantified, modelled as a random variable, and then classified as being type A or type B.

Type A uncertainties are the ones evaluated by statistical methods. Sources of uncertainty evaluated using non-statistical methods such as documented values or elicited by expertise are called type B uncertainties. For example, during a simple mass measurement, the repetition and uncertainty from calibration standards are sources of uncertainty. The former is Type A uncertainty and the latter is Type B uncertainty. However, an accurate measurement could compute the buoyancy effect and new sources of uncertainty should be taken into account such as air density changes arising from temperature or pressure variation. It is worth mentioning at this point that the measurand model plays a critical role in uncertainty evaluation.

All sources of uncertainty can be communicated using an Ishikawa diagram [21], also known as the fishbone diagram. The measurand plays the role as a principal bone and each input quantity as an adjacent bone. Each input quantity should have its uncertainty sources attached to it. The fishbone diagram gives in a single chart a qualitative summary of all sources of uncertainty for a given measurement model.

http://dx.doi.org/10.5772/66100

After quantifying the sources of uncertainty, the partial derivative of the measurand model with respect to each input quantity should be calculated. The partial derivative is called sensitivity coefficient. The standard deviation of each uncertainty source multiplied by the sensitivity coefficient gives the standard uncertainty. Considering the result obtained for a given measurand as a function of *f n* independent variables. The uncertainty can be written as:

$$u^2_{f(x,\,y,\,z,\ldots,n)} = \left(\frac{\partial f}{\partial x}\right)^2 u_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 u_y^2 + \left(\frac{\partial f}{\partial z}\right)^2 u_z^2 + \ldots, + \left(\frac{\partial f}{\partial n}\right)^2 u_n^2 \tag{1}$$

where u_f is the uncertainty associated with the measurand, the partial derivatives are the sensitivity coefficients and the terms *ui* ($i = x, y, z, \ldots, n$) are the uncertainties associated with each one of the independent variables.

The standard uncertainties could be plotted in a bar-like chart. This kind of chart is known as uncertainty budget and brings to light critical information about the measurement quality. With this chart one can easily assess the major sources of uncertainty that must be reduced in order to enhance the measurement accuracy. This last aspect points out that GUM methodology is also a management tool for continuous improvement.

In materials science measurements, the bottom-up approach is suitable for limited tests where it is possible to address the measurand via a liable model. Some cases where bottom-up approaches can be applied are the measurement of specific area by gas adsorption [22] and measurement of thermal conductivity by the guarded hot plate (GHP) method [23].

For cases where bottom-up approach does not fit or its application is much complex, the top-down approach is a suitable tool for estimating uncertainty associated with a measurand. The top-down method is a phenomenological approach that is based on the validation and the quality control tests results, assuming that these results: (i) cover all the influence factors and (ii) are representative for all measurements. In this approach, which is described in ISO 5725 [24], uncertainty will always have components evaluated from reproducibility and bias.

The top-down approach is more often used for chemical, biology and materials measurements when some relevant quantities cannot be addressed in a mathematical model. For example, the top-down method was used by De Temmerman et al. [25, 26] to evaluate the uncertainty associated with the measurement of particle size using transmission electron microscopy (TEM).

To perform the top-down approach, it is necessary to use an experimental design that addresses the method variability and a set of reference materials to estimate the bias component. This approach could be expensive and time consuming due to the large number of measurements that must be carried out in order to explore the influence of all the relevant experimental parameters on the results. One example is the use of the top-down approach to evaluate the uncertainty associated with purity analysis by using differential scanning calorimetry (DSC) [27]. Furthermore, in some cases, the applied procedure could become ineffective if future tests

must be performed under experimental conditions not contemplated in the experimental design used for uncertainty evaluation.

Besides the issues related to the use of different approaches for estimating uncertainties, measurements of material properties have some particularities that must be taken into account when evaluating their uncertainties. One of them is that these measurements are often made to give representative measurand values related to a considerable amount of material, such as a lot. In addition, in surface analysis there are some cases in which the measurement is made in a local area of a specimen instead of the material as a whole. In these cases, measured values scatter not only by reproducibility of the measurements, but also by possible non-uniformity of the material [28]. These two components are eventually merged when estimating uncertainty for a reference material. Additionally, it is worth mentioning at this point that (Section 2.4) material properties can be classified in two categories: intrinsic properties and procedural properties.

Intrinsic properties are inherent to the material and its value does not depend on the measurement procedure. On the other hand, procedural properties are totally dependent on the measurement procedure. So, in this last case, two or more property values can only be compared if the measurement procedures have been exactly the same.

3. Practical view for implementing quality management systems

3.1. Strategies for implementation and its progress monitoring

There are several ways to approach the implementation of a quality management system based on ISO/IEC 17025 in a laboratory. It is the authors' view that most of the time such a system will be implemented in a laboratory that is already in operation. Therefore, it is important that the strategy adopted for implementing the system minimizes the impact of the implementation process on the testing and R&D activities carried out by the laboratory. Furthermore, the strategy shall be flexible and prioritize constant follow-up in order to avoid repetition of work and optimize the schedule whilst minimizing the cost of the process. In this aspect, there are some points that must always be considered in the adopted strategy.

Once the organizational structure of laboratories/units has been set, as discussed in Section 2.1, the next step is to build a time chart highlighting the requirements of ISO/IEC 17025 standard which will also account for any existing internal requirements specifically applied to the provided testing services. Additionally, this time frame should specify the responsibilities of the laboratory staff members in order to organize the process. For improving the results, production and registration of technical documents related to quality assurance of testing results should be considered as the critical steps of the process. Such documents are technical standards for equipment operation, calibration and maintenance, standards for testing execution and forms used to record testing results and also any procedure related to quality assurance. Indeed, this type of implementation project management was discussed by Silva et al. [7] and, in brief, should be designed considering the following points:

http://dx.doi.org/10.5772/66100

- Clear scope definition: scope definition influences all the chronogram steps. It is crucial to clearly select the testing services that will be covered by the quality management system, and, if the laboratory is part of a large company or institute, to identify the quality management documents that dispose about laboratory activities. This item may be considered pretty basic and obvious, but these observations are essential to speed up the process;

- Connections between different requirements of ISO/IEC 17025 and other pertinent documents: the chronogram should carefully consider the connections between the different requirements of ISO/IEC 17025 and also between these requirements and other requisites from internal documents or even other standards. The understanding of the connections is important in order to avoid the superposition of tasks. It also prevents proposition of simultaneous tasks that would result better if executed in sequence;

- Impact of system implementation on on-going activities: as briefly mentioned before, it is important to consider that a quality management system is often implemented in a laboratory where some processes are already in place. When this is the case, it is crucial to minimize the impact of task implementation concerning the on-going processes. This point is especially relevant in laboratories where researchers are also responsible for the quality system management;

- Attribution of tasks: the clear designation of the staff members responsible for carrying-out each step of the project management schedule is important not only due to the practical execution of the tasks, but also to optimize the process follow-up. It also improves the communication within the team responsible for implementing the system and, consequently, the compromise between different processes that may be related; and

- Follow-up program: it is important for the chronogram to specify the follow-up strategy. The appropriate strategy depends on the case and it may work with periodic check-up of the on-going activities or with continuous monitoring.

From those points, the definition of a suitable follow-up strategy is of paramount importance for optimizing the quality system management implementation process, because it enables 'on-the-fly' changes of the adopted strategy and timeframe. When the latter is organized on the basis of requirements and all the tasks are clearly assigned, it is possible to develop a simple and efficient follow-up programme. It is easier to know how many people are assigned to specific tasks and to verify the relative progress of different tasks. This information is essential to quickly adjust chronograms and even the strategy when necessary.

A functional follow-up programme is an effective tool for checking the system's maturity determining when it is time to perform a deeper analysis in order to identify fails or gaps in the implementation process. One of the most useful tools to perform this sort of analysis is an internal auditing, which results in an improved diagnostic when compared to the previous follow-up system. This detailed diagnostic is fundamental to carry out the last adjustments necessary to finish the system implementation. The main steps of the implementation project discussed above and its relationships are shown in **Figure 2**.

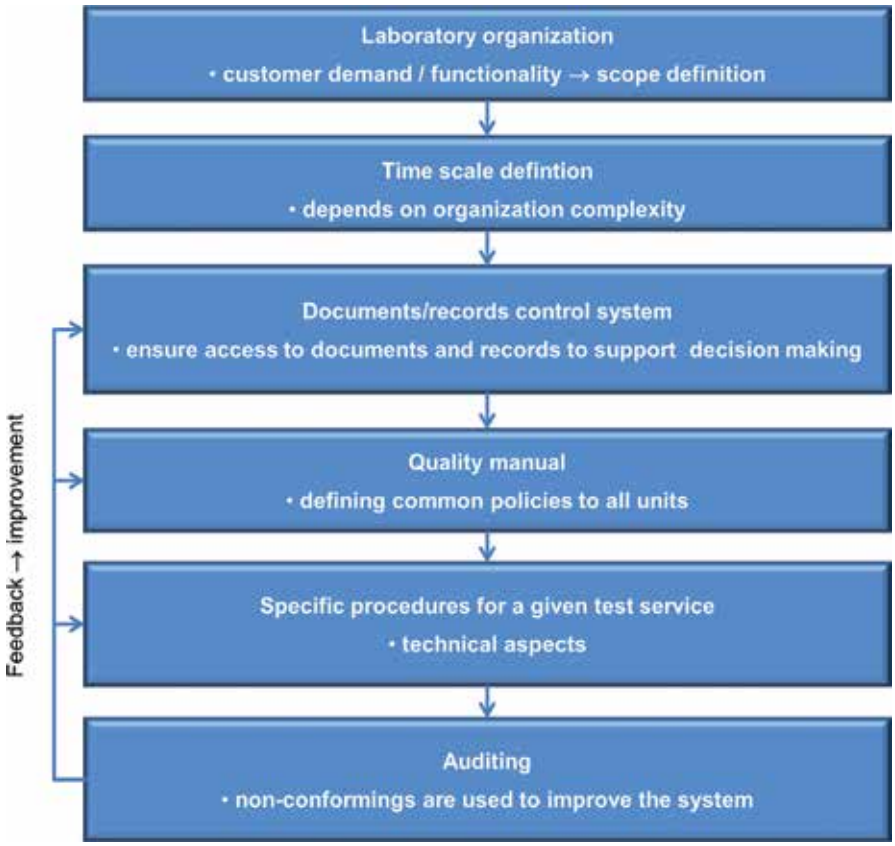

Figure 2. Main steps of a quality management system implementation project.

A last important practical aspect to be mentioned is the financial impact of implementing a quality management system at a laboratory. A deployed quality management system standardizes the processes of the laboratory, demonstrating measurement capability and technical competence. All these aspects add credibility to the results and demonstrate laboratory commitment to the customer. As a consequence, those aspects also translate into financial success by productivity increase, reduced rework and a smaller number of errors. Additionally, standardization of the supplier evaluation process, pricing policy, and careful selection of subcontracts also decreases financial costs. On the other hand, the requirements of a quality system according to ISO/IEC 17025, such as audits, calibrations, and inter-comparisons bring additional costs to the laboratory. In this aspect, Barradas and Sampaio [29], had shown that investments are feasible and customer satisfaction is improved.

3.2. Implementation time scale

The time necessary for implementing a quality management system depends basically on the size and complexity of the company/laboratory. It scales with the number of laboratories, testing services and R&D activities that the system will be applied. However, a functional

http://dx.doi.org/10.5772/66100

approach to the organization of laboratories and services decreases the time consumed for system's implementation. Indeed, the latter may be regarded as the first step thereto.

For this reason, it is somehow difficult to estimate the time scale necessary for implementing the system without considering a specific case. However, if one considers that a typical structure is composed of a number of different laboratories under a common management centre (a R&D division or department, for example), it is possible to point out some strong time scale bottlenecks:

- Definition of general principles and documents common to all the laboratories: in large structures it consists in writing a 'manual of quality', with policies and processes common to all laboratories/units under the quality management system. Such document should standardize procedures and policies (mostly those related to management requisites—item 4 in ISO/IEC17025, as for instance, control of documents and records, non-conforming work management, and management assessment), but without hindering the work at the laboratories, ensuring flexibility necessary to deal with specific characteristics of different laboratories;
- System for documents and records control: implementation of an integrated system to provide fast access to documents, to control documents update and and to keep track and security of records at a large structure composed of several laboratories may be a very time-consuming task. Ideally, such a system should also include unified policies and electronic processes for back-up that increase the implementing time;
- Metrological traceability: as earlier mentioned, ensuring metrological traceability is not always a trivial task in materials characterization. For laboratories, implementing a quality management system 'from the scratch' it may implicate both minor aspects, like acquiring certified reference materials for calibration, and very time demanding actions, like training the staff and developing validate test procedures.

4. Final considerations

The implementation of a quality management system in materials science laboratories based on the ISO/IEC 17025 standard must fulfil the requirements outlined in the standard. However, due to the particular characteristics of the work realized in such laboratories, some requirements may be considered as key matters during the system implementation' these are as follows:

- Organization of the laboratories;
- Control of documents and records;
- Customer relationship service;
- Calibration and traceability to the international system of units;

- Quality assurance of results and
- Evaluation of uncertainty of measurement.

Some of these key matters should not only fulfil the standard requirements, but also provide the necessary flexibility for the activities carried out in those laboratories, especially those related to R&D management, like control of documents and records, for example. On the other hand, some of these topics are related to technical features specific for particular tests methods for materials characterization, such as evaluation of uncertainty and traceability. These last aspects may be considered independently of managerial aspects.

The success in implementing a functional and flexible quality management system in aforementioned science laboratories depends not only on the effective and efficient handling of these key factors, but also on the use of a suitable approach for implementing the system. This approach should use appropriate strategies and chronogram in order to facilitate the implementation follow-up and minimize the time and cost of the process.

One last, but not less important, consideration should be made on further improvement of quality assurance provided by materials science laboratories. As mentioned above in this chapter, materials science impacts innovation in several areas, both on the development of new materials and on the development and optimization of techniques for materials property characterization. In this context, it is very common that the use of new materials in commercial products give rise to safety concerns about the use of the product and its disposal in the environment [30, 31]. Some typical cases are, for example, the use of nanoparticles in cosmetics and pharmaceutical products [32], and use of new composites for engineering systems [33].

These issues are not directly related to the ISO/IEC 17025, but to specific international regulation for several products and to international standards used to perform materials and products tests. In this aspect, it is important for materials laboratories engaged in quality assurance to work proactively with national and international organizations responsible for standardization and/or regulation, collaborating on the development and optimization of standard procedures for materials characterization and proposition of technical criteria for utilization of materials for specific uses.

Author details

Rodrigo S. Neves*, Daniel P. Da Silva, Carlos E. C. Galhardo, Erlon H. M. Ferreira, Rafael M. Trommer and Jailton C. Damasceno

*Address all correspondence to: rsneves@inmetro.gov.br

Materials Meteorology Division (Dimat)—National Institute of Meteorology, Quality and Technology (Inmetro), Rio de Janeiro, Brazil

References

[1] Rodima A., Vilbaste M., Saks O., Jakobson E., Koort E., Pihl V., Sooväli L., Jalukse L., Traks J., Virro K., Annuk H., Aruoja K., Floren A., Indermitte E., Jürgenson M., Kaleva P., Kepler K., Leito I. ISO 17025 quality system in a university environment. Accreditation and Quality Assurance. 2005;10:369–372. DOI: 10.1007/s00769-005-0011-x

[2] Fernandes E.A.D.N., Bacchi M.A., Tagliaferro F.S., Gonzaga C.L., De França E.J., Favaro P.C., Fogaça A.A. Quality system implementation in a Brazilian university laboratory. Accreditation and Quality Assurance. 2006;10:594–598. DOI: 10.1007/s00769-005-0061-0

[3] Zapata-García D., Llauradó M., Rauret G. Experience of implementing ISO 17025 for the accreditation of a university testing laboratory. Accreditation and Quality Assurance. 2007;12:317–322. DOI: 10.1007/s00769-007-0274-5

[4] Grochau I.H., Ferreira C.A., Ferreira J.Z., ten Caten C.S. Implementation of a quality management system in university test laboratories: a brief review and new proposals. Accreditation and Quality Assurance. 2010;15:681–689. DOI: 10.1007/s00769-010-0713-6

[5] Biasini V. Implementation of a quality management system in a public research centre. Accreditation and Quality Assurance. 2012;17:621–626. DOI: 10.1007/s00769-012-0936-9

[6] Maynard S., Foster S., Hall D. J. ISO 17025 application within racing chemistry: a case study. Technovation. 2003;23:773–780. DOI: 10.1016/S0166-4972(03)00002-6

[7] Silva D.P., Galhardo C.E.C., Lidizio L.R., Senna C.A., Damasceno J.C., Ferreira E.H.M., Trommer R.M., Fukuhara M., Neves R.S. The experience of implementing a quality management system at the Materials Metrology division (Dimat)-Inmetro: a practical approach. Accreditation and Quality Assurance. 2015;20:465–471. DOI: 10.1007/s00769-015-1156-x

[8] ISO/IEC 17025 (2005) General requirements for the competence of testing and calibration laboratories. Geneva, Switzerland.

[9] Mathur-De-Vré R. Approach to quality system in research and development. Accreditation and Quality Assurance. 1997;2:63–68.

[10] Krapp M. Quality assurance in research and development: an insoluble dilema? Fresenius Journal of Analytical Chemistry. 2001;371:704–713. DOI: 10.1007/s002160100986

[11] Baker M. Quality time. It may not be sexy, but quality assurance is becoming a crucial part of lab. 2016;529:456–458. DOI: 10.1038/529456a.

[12] ISO. ISO/IEC CD 17025 [Internet]. 09/06/2016 [Updated: 09/12/2016]. Available from: http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=66912 [Accessed: 09/19/2016].

[13] NATA. Update on the revision of ISO/IEC 17025:2005 Conformity Assessment—General requirements for the competence of testing and calibration laboratories

[Internet]. 08/2016 [Updated: 08/2016]. Available from: http://www.nata.com.au/nata/nata-enews/66-nata-e-news/oct2015/1212-update-on-the-revision-of-iso-iec-17025-2005-conformity-assessment-general-requirements-for-the-competence-of-testing-and-calibration-laboratories [Accessed: 09/19/2016]

[14] Joint Committee for Guides in Metrology. International vocabulary of metrology–basic and general concepts and associated terms (VIM). 3rd ed. JCGM/BIPM; 2012. 91 p.

[15] Salas-Tellez J.A., Guardado-Perez J.A., Rosas-Gutierrez F., Mitani Y. A traceability scheme for materials metrology. Metrologia. 2010;47:S18–S22. DOI: 10.1088/0026-1394/47/2/S02

[16] Roebben G., Insinger T., Lamberty A., Emons H. Metrological traceability of the measured values of properties of engineering materials. Metrologia. 2010;47:S23–S31. DOI: 10.1088/0026-1394/47/2/S03

[17] Bureau International des Poids et Mesures. The BIPM key comparison database [Internet]. Available from: http://kcdb.bipm.org/ [Accessed: 29/09/2016]

[18] Versailles Project on Advanced Materials and Standards (VAMAS)—Formation and Objectives [Internet]. Available from: http://www.vamas.org/formation.html [Accessed: 07/29/]

[19] NIST. Certificate—SRM 676a—Alumina Powder (Quantitative Analysis Powder Diffraction Standard) [Internet]. [Updated: 3/16/2016]. Available from: https://www-s.nist.gov/srmors/view_cert.cfm?srm=676A [Accessed: 09/01/2016]

[20] JCGM 100:2008 Evaluation of measurement data—Guide to the expression of uncertainty in measurement. 1st ed. BIPM/JCGM; 2008. 120 p.

[21] Kindlarski E. Ishikawa Diagrams for problem solving. Quality Progress. 1984;17:26–30.

[22] De Lange M.F., Vlugt T.J.H., Gascon, J., Kapteijn F. Adsorptive characterization of porous solids: error analysis guides the way. Microporous and Mesoporous Materials. 2014;200:199–215. DOI: 10.1016/j.micromeso.2014.08.048

[23] NIST. Assessment of Uncertainties for the NIST 1016 mm Guarded-Hot-Plate Apparatus: Extended Analysis for Low-Density Fibrous-Glass Thermal Insulation. [Internet]. 2009. Available from: http://fire.nist.gov/bfrlpubs/build09/art003.html [Accessed: 09/01/2016]

[24] ISO 5725-1 Accuracy (trueness and precision) of measurement methods and results—Part 1: General principles and definitions. ISO; 1994.

[25] De Temmerman P.-J., Lammertyn J., De Ketelaere B., Kestens V., Roebben, G., Verleysen E., Mast J. Measurement uncertainties of size, shape, and surface measurements using transmission electron microscopy of near-monodisperse, near-spherical nanoparticles. Journal of Nanoparticles Research. 2014;16:2177. DOI: 10.1007/s11051-013-2177-1

[26] De Temmerman P.-J., Verleysen E., Lammertyn J., Mast J. Size measurement uncertainties of near-monodisperse, near-spherical nanoparticles using transmission electron

microscopy and particle-tracking analysis. Journal of Nanoparticles Research. 2014;16:2628. DOI: 10.1007/s11051-014-2628-3

[27] Kestens V., Roebben G., Linsinger T. Development and validation of a differential scanning calorimetry purity determination method for polycyclic aromatic hydrocarbons. Accreditation and Quality Assurance. 2010;15:269–281. DOI: 10.1007/s00769-010-0632-6

[28] Baba T. Measurements and data of thermophysical properties traceable to a metrology standard. Metrologia. 2010;47:S143–S145. DOI: 10.1088/0026-1394/47/2/S12

[29] Barradas J., Sampaio P. ISO 9001 and ISO 17025 standards in a metrology laboratory. In: 14th Toulon-Verona Conference "Organizational Excellence in Services"; September 1–2 2011; Alicante/Spain. Alicante/Spain: University of Alicante; 2011. pp. 143–152.

[30] Wiesner M.R., Lowry G.V., Alvarez P., Dionysiou D., Biswas P.Assessing the risks of manufactured nanomaterials. Environmental Science & Technology. 2006;40(14):4336–4345.

[31] Marano F., Guadagnini R. Health impacts of nanomaterials. In: Lourtioz J.-M., Lahmani M., Dupas-Haeberlin C., Hesto P., editors. Nanosciences and nanotechnology: Evolution or revolution? 1st ed. London: Springer International Publishing; 2016. pp. 273–286. DOI: 10.1007/978-3-319-19360-1_12

[32] Wacker M.G., Proykova A., Santos G.M.L. Dealing with nanosafety around the globe—Regulation vs. innovation. International Journal of Pharmaceutics. 2016;509(1–2):95–106. DOI: 10.1016/j.ijpharm.2016.05.015

[33] Denning S. What went wrong at Boeing. Strategy and Leadership. 2013;41(3):36–41. DOI: 10.1108/10878571311323208

Chapter 3

Youden Two-Sample Method

Julia Martín, Nieves Velázquez and
Agustin G. Asuero

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66411

Abstract

The results obtained when testing materials, equipment and procedures are not generally identical. Factors that influence the magnitude of the results are not fully controllable. As such, the interpretation and analysis of results must take into account the variations caused by numerous and random unavoidable causes. Intercomparison exercises are considered of being of importance, as they do allow the examination of the analytical process and their generated results. Youden plot is particularly aimed at interlaboratory comparisons. The raw results provided by the participating laboratories are treated by a statistical method applied by the centre performing the trial. In order to materialize this, two similar materials with small differences in the concentration of the characteristics are required. The advantage of Youden analysis is its ability to separate the random errors with a minimum effort by participants in the design from the point of view of the analytical requirement. This book chapter illustrates the method that has been applied to elaborate on data covering a diverse scientific field: polyunsaturated fatty acids in fat and oils, total blood cholesterol and aspirin in pharmaceutical preparations. Finally, liquid chromatography with tandem mass spectrometry detector has been applied to the determination of an emerging contaminant, methylparaben (MeP), in surface waters.

Keywords: Youden plot, confidence ellipse, quality control

1. Introduction

The main objective of quality systems when implanted in analytical laboratories is to ensure that the results obtained confirm to quality standards, in addition to them showing a level of harmonization [1–4] between obtained results.

In order to achieve this goal, quality assessment systems are implemented so as to allow the examination of the analytical process as well as of their results generated.

Quality assessment is coined as the systematic examination carried out by an entity to verify [5–7] that it meets specified requirements (fitness for purpose). This is a generic concept that can be refined more by relating it to the specific set of activities planned and executed with the aim of ensuring that the activities involved in the quality control are done in a proper and efficient way.

The quality assessment involves the methodical and continuous contrast of the product, system or quality service. In the specific area of the laboratory, it refers to examination of systems and to analytical results generated both in terms of accuracy [8, 9] and representativeness.

Intercomparison exercises are framed in this context [10–13], establishing the procedure for design, organizing and gathering information from a set of laboratories working with the same samples that undergo an assessment of their results. An intercomparison exercise is based on acceptance by several laboratories to perform the same analysis. This analysis is carried out under the co-ordination of an organization. The purpose is to assess the quality of their work, to evaluate the method of measuring, to determine the property of a material (the content of an element or compound, etc.).

The main mission of the organization is to establish the objectives and the conditions regarding the participation of the laboratories [14–18] while ensuring the quality and stability of the sample under study. Those institutions are also responsible of dealing with the statistical treatment of obtained outcomes. The participating laboratories should, in turn, commit themselves to follow the conditions set by the organization, which may change depending on the type of executed exercise.

A very important aspect of intercomparison exercises is the selection of material to be used for the study. In that sense the type of matrix and the nature and range of values of the parameter (or parameters) under study must be defined.

It is worth mentioning at this point that not all materials are suitable to carry out a study of this type. It is essential that the material used is representative, homogeneous and stable. The organization is responsible for ensuring that the criteria mentioned above are met.

The preparation of the material must follow a series of stages, after which the material is packaged. It should, hence, be homogeneous and stable. The submitted sample must be properly identified and packaged in order to prevent breakage.

The sample is tagged accordingly so as to show the state in which the sample has been sent to the participating laboratory. Guidelines for sample preservation and handling, a description of the analytical methods to be applied as well as the report methods must also be included in the shipment.

Youden two-sample diagram, two-sample collaborative testing, two-sample plan, Youden plot or Youden analysis is particularly aimed at interlaboratory comparisons [19, 20]. The raw results provided by the participating laboratories are treated by a statistical method applied by the organizing centre of the trial. The Youden approach or z-score marks are some of the tools used for the treatment of the results. Finally, based on these statistical treatments, a statement is sent to the laboratories that have presented inaccurate results, as well as appropriate suggestions in order to improve their work.

A literature search has been carried out in order to validate current status with regards to the use of the Youden approach. The information gathered is presented in **Table 1** [1–129].

Performed literature review reveals that the Youden chart has been successfully used in agriculture, environmental chemistry, geochemistry, industry and medicine. The invention of the Youden diagram may be regarded as setting a landmark in quality control in clinical chemistry [21].

The performance and evaluation of interlaboratory programmes by Youden's method are suited to laboratory monitoring and allow to obtain information concerning both precision and systematic errors from analytical results without much effort.

Content	Reference
Key feature in analytical proficiency testing	[5]
Comparative studies of Shewhart, Thompson, Howarth and Youden representations: advantages and disadvantages	[21]
Implementation of two new graphical methods recommended by the ISO standard, Mandel's h statistic and the Youden plot, to evaluate the consistency between laboratories and within laboratories for radon and thoron exposures	[22]
A proficiency testing scheme (CNAS T0419) is described involving 217 laboratories in China as participants using their regular analytical methods for the determination of lead and arsenic in foundation cream cosmetics	[20]
An optimized Youden chart was developed and compared with the traditional and trimmed traditional Youden charts	[23]
A robust Youden plot is constructed based on robust statistical parameters since these are scarcely affected by non-normally distributed data, and this approach is applied in an external quality assessment (EQA) programme.	[24]
Youden representation for mycotoxins (deoxynivalenol and ochratoxin A) and toxins (T-2 and HT-2) in wheat and corn	[25]
Metrology statistical manual: the Youden approach with standardized variables	[26]
Interlaboratory studies: statistical organization protocol and evaluation	[27]
Control blood for an external quality assessment scheme (EQAS) for international normalized ratio (INR) point-of-care testing (POCT) in the Netherlands and to assess the performance of the participants	[28]
Application of ISO 13528 robust statistical methods for external quality assessment of blood glucose measurements in China	[29]
A study under what conditions of measurement to assess bias and from the results of a six-round blind-duplicated interlaboratory proficiency programme for creatinine in urine shows that bias is present in each individual run with components from that batch and from the laboratory over the rounds of the programme	[30]
Brazilian interlaboratory programme study on anion measurement in synthetic water. The programme described is promoted regularly since 2007 and recommended the use of ion chromatography as analytical technique for all participant laboratories	[31]
Robust determination of the correlation coefficient, analytically validated using two types of statistical models and computational simulations	[32]
Comparison of the statistical Youden method (by Hotelling T2 test and bivariate normal distribution) in interlaboratory studies by ISO and National Association of Testing Authorities (NATA) standards	[33]
Collaborative study procedures	[14]
A method validation study was conducted according to the IUPAC harmonized protocol for the determination of ochratoxin A in *Capsicum* spp. (paprika and chilli). The study involved 21 participants representing a cross section of research, private and official control laboratories from 14 EU member states and Singapore	[34]

Content	Reference
Collaborative test on the count of *Escherichia coli*	[35]
Proficiency test for the determination of heavy metals in mineral feed. The importance of correctly selecting the certified reference materials during method validation	[36]
Evaluation tools to understand statistical methods related to the *z*-score for use in proficiency testing by interlaboratory comparisons	[37]
Evaluation of learning outcomes in quantitative analysis lab using Youden plots	[38]
State of the art with respect to the selection and use of proficiency testing schemes and the interpretation of results and evaluations given in proficiency testing schemes	[39]
Performances of analytical methods for atmospheric deposition and soil analysis assessed through intercomparison exercises	[40]
HIV external quality assessment (EQA) results by the KCDC from the 17 HIV testing laboratories that also performed HIV-1 western blot testing of the 585 laboratories	[41]
Second interlaboratory exercise on non-steroidal anti-inflammatory drug analysis in environmental aqueous samples	[42]
Statics and chemometrics for analytical chemistry	[7]
A proficiency testing scheme was developed for a limited number of analytical laboratories participating in the analysis of natural water in Israel	[43]
Proficiency test for heavy metals in feed and food in Europe	[44]
A multilaboratory proficiency testing programme was conducted by the National Accreditation Board for Testing and Calibration Laboratories (India) and coordinated by the Institute of Pesticide Formulation Technology. This programme was conducted to compare the performance of individual laboratories in the area of pesticide formulation (Chlorpyrifos 20 EC) analysis. A total of 24 laboratories in India participated	[45]
Proficiency testing for the determination of pesticides in mango pulp: a view of the employed chromatographic techniques and the evaluation of laboratories' performance	[46]
Investigations for the improvement of the measurement of volatile organic compounds from floor coverings within the health-related evaluation of construction products: application of the Youden method	[47]
Characterization of candidate reference materials for bone lead via interlaboratory study and double isotope dilution mass spectrometry	[48]
Implementation and methodology of an interlaboratory system that ensures the quality of glassware calibration and use in a large laboratory	[49]
An updated liquid chromatographic assay for the determination of glyphosate in technical material and formulations: application of the Youden method	[50]
Collaborative studies for quantitative chemical analytical methods	[51]
Description and results of the 2005 interlaboratory comparison exercise for trace elements in marine mammals. Two quality control materials derived from fresh-frozen marine mammal livers were produced and characterized at the NIST and were then distributed to over 30 laboratories	[52]
Youden method applied to the external quality control of semen analysis in Germany	[53]
Quality assurance in analytical chemistry application in the environmental, food and materials analysis, biotechnology and medical engineering	[1]
Brief note on the Youden method	[54]
Interlaboratory comparison by means of method performance precision and bias studies and proficiency testing schemes are described. The set-up of the experiments and the evaluation of the data by means of graphical and statistical methods are considered	[11]

Content	Reference
Practical advice on the Youden plot	[55]
Repeatability and reproducibility of determination of the nitrogen content of fishmeal by the combustion (Dumas) method and comparison with the Kjeldahl method: interlaboratory study	[56]
Application of the Youden method in clinical chemistry: cortisol determination	[57]
An investigation of the capability of the medium resolution imaging spectrometer validation teams to determine chlorophyll a, using the latest measuring protocols and advanced high-performance liquid chromatography and spectrophotometric and fluorometric method has been performed	[58]
Standardization of calibration and quality control surface-enhanced laser desorption/ionization time of flight mass spectrometry	[59]
A chlorophyll-a interlaboratory comparison was carried out to compare three different analytical chlorophyll-a determination methods: a German standard DIN 38412-16, a method of the HELCOM-Combine-Manual and the different "in-house" methods of participating laboratories	[60]
Results for total chloride content in four different types of Portland cement provided by testing laboratories participating in an interlaboratory comparison are presented. The data sets were evaluated by using different statistical methods	[61]
A proficiency test on the quantification of trace elements in serum was carried out to verify the performance of about 30 regional laboratories of the network of Italian laboratories. The exercise consisted of four runs in which the laboratories were free in choosing analytical methods to determine trace elements in freeze-dried animal serum. Laboratory performances were evaluated by the study of statistical functions as coefficients of variation (CV), Youden plot and z-score value	[62]
Collaborative studies for cereal analysis	[10]
Practical digest for evaluating the uncertainty of analytical assays from validation data according to the LGC/VAM protocol	[63]
Statistical methods for use in proficiency testing by interlaboratory studies, International Organization for Standardization	[13]
Youden analysis of Karl Fisher titration data from an interlaboratory study determining water in animal feed, grain and forage	[64]
Establishing measurement traceability in clinical chemistry: cholesterol, progesterone and aldosterone in serum	[65]
Worldwide and regional intercomparison for the determination of organochlorine compounds and petroleum hydrocarbons in mussel tissue IAEA-432	[66]
Interlaboratory study on the determination of ascorbic acid in serum	[67]
An intercomparison of in vitro chlorophyll-a determination	[68]
Guide about collaborative studies to validate characteristics of an analytical method	[69]
Youden method application to result in total and dissolved organic carbon in surface waters	[70]
Interlaboratory exercise conducted within the framework of a hydrological project on underground water	[71]
Interlaboratory study on the determination of trace elements in sea water	[72]
State of the art with respect to the selection and use of proficiency testing schemes and the interpretation of results and evaluations given in proficiency testing schemes	[15]
Interlaboratory studies in analytical chemistry: method performance studies (collaborative trials), laboratory-performance studies (proficiency tests), collaborative bias evaluation, interlaboratory evaluation of to-be standard methods as well as certification studies for reference materials	[12]

Content	Reference
Intercomparison exercise on the determination of organochlorine compounds and petroleum hydrocarbons in algae	[73]
Statistical model assumptions upon which the procedure is based. Provides validity tests for several of these assumptions, explains conditions under which Youden is not consistent with precision estimate and indicates when precision estimates based on the procedure should be interpreted with caution or should not be used	[74]
Intralaboratory testing of method accuracy from recovery assays	[8]
Application of the Youden method to the mass fraction Youden protein fodder	[75]
Succinct description of the two-sample Youden method	[19]
Performances of analytical methods for freshwater analysis assessed through intercomparison exercises	[76]
Proposed guidelines for the internal quality control of analytical results inthe medical laboratories	[77]
Application and improvement of the Youden analysis in the intercomparison between flowmeter calibration facilities	[78]
Basic of interlaboratory studies: the trends in the new ISO 5725 standard edition	[79]
Round-robin study of performance evaluation soils vapor-fortified with volatile organic compounds	[80]
Protocol for the design, conducting and interpretation of collaborative studies	[2]
Application of the Youden method to acid rain analites	[81]
A bivariate control chart for paired measurements	[82]
Polystyrene film as a standard for testing FT-IR spectrometers	[83]
Graphical diagnosis of interlaboratory quality control data for surface water samples	[84]
Nomenclature of interlaboratory analytical studies	[85]
Basic method for the determination of repeatability and reproducibility of a standard measurement method	[86]
Reviews on the life and work of Youden	[87]
Quality control in analytical chemistry	[6]
Assessment of overall accuracy of lead isotope ratios determined by inductively coupled plasma mass spectrometry using batch quality control and the Youden two-sample method	[9]
World Health Organization international intercalibration study on dioxins and furans in human milk and blood	[88]
Analytical quality assurance. A review	[89]
Multiway analysis of variance for the interpretation of interlaboratory studies	[90]
External quality control study on the reliability of current histamine determinations in European laboratories	[91]
Guidelines for the development of standard methods of collaborative study: organization of interlaboratory studies and a simplified approach to the statistical analysis of collaborative study results	[16]
Classic paper reprint. The collaborative test of Youden	[92]
Robust statistic and functional relationship estimation for comparing the bias of analytical procedures over extended concentration ranges	[93]
Bias-free adjustment of analytical methods to laboratory samples in routine analytical procedures	[94]
Protocol for the design, conducting and interpretation of method-performance studies	[3]
Exchange of comments on a new technique in chemical assay calculations	[95]

Content	Reference
Measurement, statistics and computation, analytical chemistry by open learning. Application to aspirin preparations	[96]
Quality assurance of chemical measurements	[4]
The use of statistics to develop and evaluate analytical methods	[17]
Interlaboratory evaluation of high-performance liquid chromatographic. Determination of nitroorganics in munition plant wastewater	[97]
Interlaboratory variability in trace element analysis	[98]
The limitations of models and measurements as revealed through chemometrics intercomparison	[99]
Considerations about the graphical representation	[100]
Reverse-phase HPLC method for analysis of TNT, RDX, HMX and 2,4-DNT in munitions wastewater	[101]
Determination of heavy metals in reference marine sediments. Application of the Youden method	[102]
Organization and evaluation of interlaboratory comparison studies amongst southern African water analysis laboratories	[103]
The use of the Youden plot for internal quality control in the immunoassay laboratory	[104]
An annotation on the Youden method: recognition of the systematic and random errors	[105]
Testing laboratory performance: evaluation and accreditation	[106]
Qualification of estimates for total trace elements in food stuffs using measurement by atomic-absorption spectrophotometry	[107]
A collaborative study for measuring polyunsaturated fatty acids in fats and oils	[108]
Application of interlaboratory studies on the quality of effluent wastewaters	[109]
Statistical techniques for collaborative tests. Planning and analysis of results of collaborative tests	[18]
Interpretation and generalization of Youden's two-sample method	[110]
Collaborative analysis and the standardization of analytical methods	[111]
Graphical diagnosis of interlaboratory test results (reprinted from industrial quality control)	[112]
Systematic versus random error laboratory surveys	[113]
Precision measurement and calibration. Statistical concepts and procedure	[114]
A graphic display of interlaboratory test results	[115]
Determination of systematic an accidental errors of analytical procedure by the Youden method	[116]
Collaborative test	[117]
The sample, the procedure and the laboratory	[118]
Graphical diagnosis of interlaboratory test results	[119]
Statistical aspects of the cement testing programme	[120]
Evaluation of chemical analyses on two rocks. A simple graphical technique is proposed to aid in the comparisons between laboratories	[121]
A plan for studying the accuracy and precision of an analytical procedure	[122]
Design and interpretation of interlaboratory studies of test methods	[123]

Table 1. Some published papers dealing with the Youden approach.

2. Literature review: the Youden plot

W. J. Youden (1900–1973) was a physical chemist during the first third of this life, who turned into a statistician later, employed by the National Bureau of Standards (NBS) (now National Institute of Standards and Technology, NIST) from 1948 until his death in 1971. One of his more memorable sentences states [22, p. 12] that "The best way to find out about some of the difficulties in making measurements is to make measurements" [22].

He approached interlaboratory testing as a means of uncovering biases in measurement processes, and the so-called Youden plot has become an accepted design and analysis technique throughout the world for comparing precision and bias amongst laboratories. Youden suggested in 1959 a very simple graphical procedure for plotting results obtained by different laboratories [23–25]. Work in graphical methods, which began with the Youden plot, continues today, notably in recent works of NIST chemists.

The above is also referred to as two-sample collaborative testing, two-sample diagram, two-sample plan or Youden plot.

The method focuses on intercomparison exercises. The main characteristic is its ability to separate the systematic and random errors with minimal effort on the part of the participants.

The method is implemented as follows:

Two nearly identical samples are prepared, divided and sent to each of the participating laboratories, as recommended by Youden. A scatter plot is drawn in which the x-axis indicates one of the reported values and the y-axis the other. The scale units are the same along each axis. Each pair of results, corresponding to a given laboratory, is a point in the Youden plot (see **Figure 1**).

The points will cluster in a circular pattern whose centre is the mean values for the two samples.

Once the results are represented in the plot, they are divided into four quadrants, which are identified as (+, +), (−, +), (−, −) and (+, −). When any laboratory's result exceeds the mean achieved for all laboratories, a plus sign is used, a minus sign indicates a value smaller than the mean. If the variation in results is dominated by random errors, it would be expected that the points fall randomly distributed in all quadrants, with similar number of points in each quadrant. When systematic errors are significantly larger than random errors, then the points occur primarily in the (+, +) and the (−, −) quadrants, forming an elliptical pattern around a line bisecting these quadrants at a 45° angle.

The plot is an effective method to qualitatively evaluate the results and the capabilities of the proposed method. As can be seen in **Figure 2**, the length of a perpendicular line from any point to the 45° line is proportional to the contribution of random error on a given laboratory's results (red arrow). The distance from the intersection of the axes (mean values for samples X and Y) to the perpendicular projection of a point on the 45° line is proportional to the laboratory's systematic error (green arrow).

http://dx.doi.org/10.5772/66411

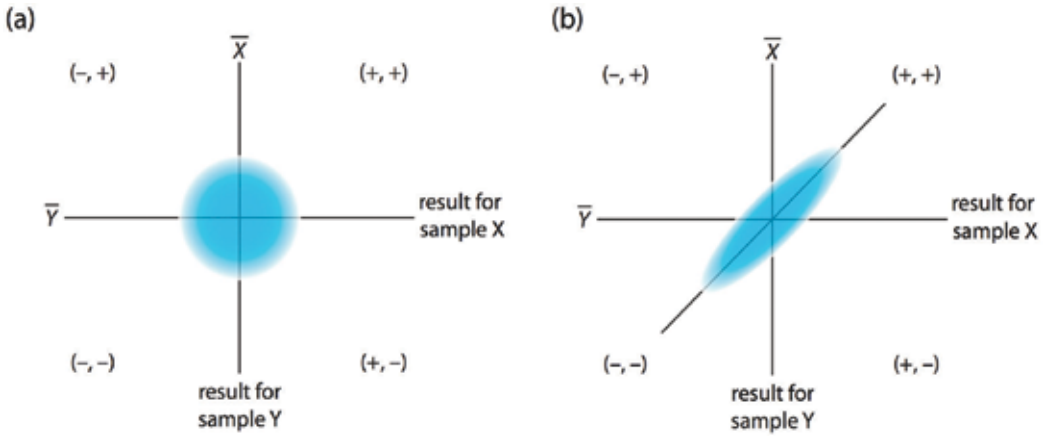

Figure 1. Typical Youden plots when (a) random errors are significantly larger than systematic errors due to the analysts and (b) when systematic errors due to the analysts are significantly larger than the random errors [126].

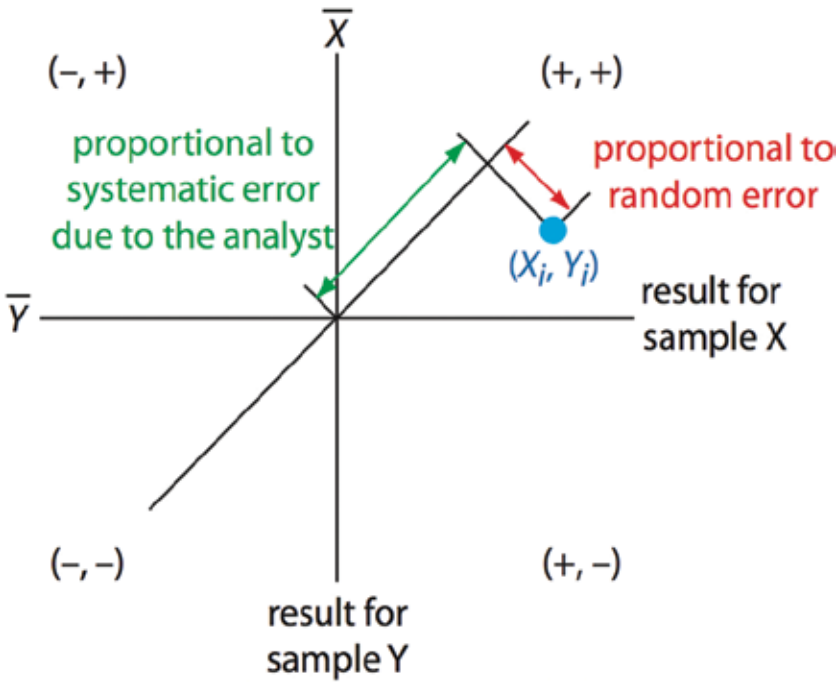

Figure 2. Relationship between the result for a single laboratory (in blue) and the contribution of random error (red arrow) and the contribution from the laboratory's systematic error (green arrow) [126].

An ideal standard method is linked with small random and systematic errors characterizing a circular compact cluster of points.

The Youden plot is a special case of the bivariate control chart to evaluate the performance of several laboratories, and the idea behind is the principal component analysis [26].

In 1974, the Youden plot was extended by Mandel and Lashof by using an ellipse instead of a circle [27]. In Youden's original method, the concentration of the analyte in the two materials was nearly the same, so that the repeatability as well as the laboratory biases would be the same for two materials [28].

Mandel and Lashof investigate the situation where two samples do not have a similar concentration so that random and systematic errors are no longer necessarily the same for both methods. They showed that in all cases, the points in the plot fall within an elongated ellipse. When Youden's original plot (two similar samples) is applied to, then the major axis forms a 45° angle. In retrospect, when samples are not similar to one another, different angles may be obtained. Their paper contains a procedure to decide whether lab bias occurs or not and also contains an estimate of all variance components.

The confidence ellipse has been proposed in ISO 13528:2005 to indicate anomalies in the between- and the within-laboratory errors in qualitative terms. For laboratory monitoring, interlaboratory tests performed according to ISO 5725-2 require much effort, especially because a large volume of samples must be provided by the organizer for $K = 4$ repeated analyses per laboratory. It is worth noting at this point that this is only suited for process standardization.

As already mentioned (see pp. 3–4), the performance and evaluation of interlaboratory programmes by Youden's method are recommended above all for laboratory monitoring. It allows to obtain information concerning both precision and systematic errors from analytical results without much effort. In addition to the above, Youden's method requires less effort for organizers and participants alike. It equally showcases a simple evaluation meaning that a potential manipulation is less likely.

Youden's method has been recommended in modern statistical manuals, procedures and protocols [29, 30] as well as recent papers, reports and government agencies, as depicted in **Table 1** [31–34].

This study discusses the Youden method and elaborates its applications in a number of diverse areas.

The experimental systems selected first for gaining experience and training in the application of the method have been the determination of polyunsaturated fatty acids in fats and oils [35], total blood cholesterol [36] and aspirin in pharmaceutical preparations [37], i.e. food and clinical and pharmaceutical applications, respectively.

Finally, a detailed procedure for the determination of methylparaben in surface waters, of special relevance nowadays in the environmental field, has been developed by using liquid chromatography-mass spectrometry.

At last, the confidence ellipse as proposed by ISO 13528:2005 will be described, and a practical case of concentrations of antibodies for two similar allergens is used as an aid to interpret the plot.

http://dx.doi.org/10.5772/66411

2.1. Youden plot development

The Youden plot is developed as follows:

1. Draw on the graph the points (X, Y) with the results submitted by the laboratories and reject any obvious anomalous points.
2. Calculate the centroid (X-mean, Y-mean) and draw up the lines. The vertical line is the average value for sample X, and the average value for sample Y is shown by the horizontal line.
3. Draw up the line X = Y passing through the centroid.
4. The difference, D, between the results $D_i = X_i - Y_i$ is referred to as random error. To estimate the total contribution from random error, the standard deviation of these differences, S_D, for all laboratories is used as follows:

$$S_D^2 = \frac{\Sigma (Di - \text{Dmean})^2}{2(l-1)} \tag{1}$$

where l is the number of laboratories and the factor of 2 is the result of using two values to determine D_i.

5. In the same way, the total, T, of each laboratory's results ($T_i = X_i + Y_i$) contains contributions from both random error and twice the laboratory's systematic error:

$$\sigma_{tot}^2 = \sigma_{rand}^2 + 2\,\sigma_{syst}^2 \tag{2}$$

6. The standard deviation of the totals, S_T, provides an estimate for σ_{tot}:

$$S_T^2 = \frac{\Sigma(Ti - \text{Tmean})^2}{2(l-1)} \tag{3}$$

Again, the factor of 2 in the denominator is the result of using two values to determine T_i.

7. If the systematic errors are significantly larger than the random errors, then S_T is larger than S_D, a hypothesis that can be evaluated using a one-tailed F-test, where the degrees of freedom for both the numerator and the denominator are n−1.
8. If S_T is significantly larger than S_D, σ_{tot}^2 may be split into components representing random error and systematic error:

$$\sigma_{SIS}^2 = \frac{S_T^2 - S_D^2}{2} \tag{4}$$

9. Calculating the radius of the confidence circle:

$$R = SD \cdot b \tag{5}$$

$$b = -2\ln\left(1 - \frac{P}{100}\right) \tag{6}$$

$$P = 100\left(1 - \exp\left(\frac{-b^2}{2}\right)\right) \tag{7}$$

where % P is the percentage of selected confidence level (usually 95%).

10. Draw a circle with radius R and centroid (X-mean, Y-mean). The laboratories falling outside the 95% circle are said to provide biased results. The radius of the circle is based on a multiple of SD, depending on the desired percentage of observations anticipated to fall within a bivariate normal distribution. A circle whose radius is a multiple of SD = (2.5; 3) represents the smallest circle that can be contained in almost every point, in the absence of bias.

2.2. An alternative approximation: the *Z*-score

An alternative to Youden Plot is the punctuation *Z*-Score. This value is used to "score" a parameter in a particular round of a laboratory's participation. This is done by means of the following calculations:

1. Calculate the median of X and Y.
2. Calculate the total error (ε_T) for each laboratory:

$$\varepsilon_T = \sqrt{(x_i - x_{\mathrm{Me}})^2 + (y_i - y_{\mathrm{Me}})^2} \tag{8}$$

3. Calculate the systematic error component (C_s) as

$$C_s = \frac{(x_i - x_{Med}) + (y_i - y_{\mathrm{Me}})}{\sqrt{2}} \tag{9}$$

4. Calculate the random error component (C_R) as

$$C_R = \sqrt{\left(x_1 - \left(\frac{C_S}{\sqrt{2}}\right) + x_{\mathrm{Me}}\right)^2 + \left(y_1 - \left(\frac{C_S}{\sqrt{2}}\right) + y_{\mathrm{Me}}\right)^2} \tag{10}$$

5. Systematic and random components calculation required so that their sum is equal to the magnitude of the total error:

Systematic error:

$$\varepsilon_S = \frac{C_s}{|C_S| + C_R}\varepsilon_T \tag{11}$$

Random error:

$$\varepsilon_R = \frac{C_R}{|C_S| + C_R}\varepsilon_T \tag{12}$$

Thus,

$$E_T = |\varepsilon_S| + \varepsilon_R \tag{13}$$

6. Calculate the typical deviation:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \varepsilon_{Ri}^{2}}{n-1}} \tag{14}$$

where n is the number of laboratories.

7. Finally, the z-score is calculated as

$$z = \frac{x_i - x_{\mathrm{Me}}}{\sigma} \tag{15}$$

The results are classified as

$\lvert z \rvert \leq 2$	"satisfactory"
$2 < \lvert z \rvert \leq 3$	"questionable"
$\lvert z \rvert > 3$	"unsatisfactory"

3. Application to experimental systems

3.1. Determination of polyunsaturated fatty acids in fats and oils

To illustrate the procedure, data from the interlaboratory study of a method for determining polyunsaturated fatty acids (PUFA) in fats and oils is used. The procedure consisted of saponifying a sample, treating it with an enzyme and measuring the absorbing product at 234 nm. Palm oil, corn oil and three hydrogenated blends were used in the study. One blend of hydrogenated oil was separated into two parts and designated as samples X and Y, respectively. Random subsamples from the two samples were analysed in each of the 17 laboratories as blind duplicates. The test incorporated an official Food and Drug Administration (FDA) method and a slightly modified one using boron trifluoride-methanol (BF). The aim in this case was to identify the laboratories that have higher quality results.

The results for FDA method are shown in **Table 2**. The data from laboratory 1 were rejected because inaccurate results were proportioned; those from laboratory 11 are listed but were not used (because their results were beyond the ones shown by others (8.2 g/100 g for sample X and 26.3 g/100 g for sample Y) in the calculations.

The vertical line at 28.6 g/100 g is the average value for sample X, and the average value for sample Y is shown by the horizontal line at 28.2 g/100 g. To estimate σ_{rand} and σ_{syst}, the values for D_i and T_i are calculated first. Next, the standard deviations for the differences, S_D, and the totals, S_T, are computed using Eqs. (1) and (3), yielding $S_D = 1.53$ and $S_T = 3.11$. To determine if the systematic errors between the laboratories are significant, the F-test is applied so as to compare S_T and S_D.

Because the F-ratio (4.141) is larger than F(0.05,14,14), which is 2.484, it is concluded that the systematic errors between the analysts are significant at the 95% confidence level, which is estimated using Eq. (4) giving 3.67.

Laboratory	FDA method		Laboratory	BF method	
	Sample X	Sample Y		Sample X	Sample Y
2	26.1	28.5	2	24.9	29.4
3	29.6	28.6	3	30.3	29.4
4	29.2	26.8	4	29.1	31.7
5	29.5	26.9	5	31.4	29.3
6	30.3	30.8	6	29.1	30.4
7	27.5	25.9	7	26.6	26.6
8	25.8	26.9	8	30.0	30.7
9	30.0	28.0	9	29.5	29.7
10	29.0	25.0	10	28.3	29.3
11*	8.20	26.3	11*	10.3	29.6
12	31.3	32.0	12*	10.5	10.1
13	24.7	24.8	13	25.3	24.8
14	24.3	25.9	14	26.3	28.6
15	31.0	31.3	15	31.4	30.3
16	28.2	32.3	16	28.0	28.0
17	31.8	29.9	17	29.6	27.0

*Not included in mean.

Table 2. Determination of cis, cis-PUFA in blind duplicate samples by two methods (g trilinolein/100 g sample).

The results are plotted in **Figure 3**. The latter reveals that laboratories 11 (aberrant result), 12, 13, 14, 15 and 16 are outside the 95% circle, indicating high systematic errors.

The results for the BF method are shown in **Table 2**.

Again, the data from laboratory 1 were rejected because of a mistake; those from laboratories 11 and 12 are listed but were not used (10.3 and 10.5 g/100 g, respectively, for sample X and 29.6 and 10.1 g/100 g, respectively, for sample Y), as they lie beyond the set limit [35] in the calculations.

Again, the F-ratio (3.268) is larger than F(0.05,13,13), which is 2.577, so it is determined that the systematic errors between the laboratories are significant at the 95% confidence level. All the results are plotted in **Figure 4**. By observing the latter, one may observe that the laboratories 11, 12 (aberrants) and 2 and 13 are outside the 95% circle and are displaced far from the cluster of the others.

Notice that in both methods, about half the points lie above, and about half lie below the horizontal lines through the two means. Likewise, the vertical lines also separate the laboratories into equal groups, as do the 45° lines. However, in neither plot are the results equally distributed amongst the four quadrants; there are more in the upper right and lower left quadrants than in the upper left and lower right. Dispersion along the 45° line indicates that laboratories are high or low on both samples, while dispersion at right angles to the 45° line indicates a lack of agreement between results from the same laboratory.

http://dx.doi.org/10.5772/66411

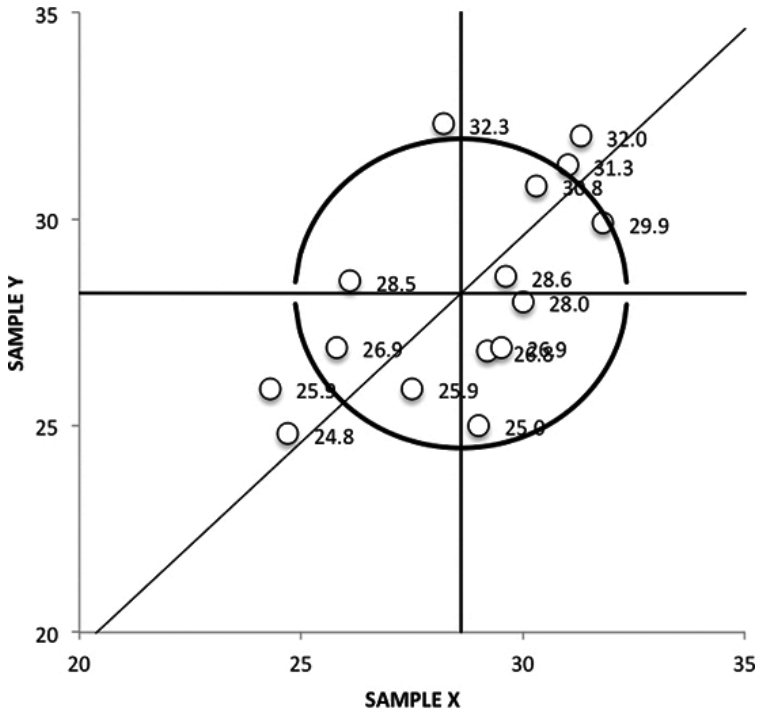

Figure 3. Youden plot (determination of PUFA, FDA method).

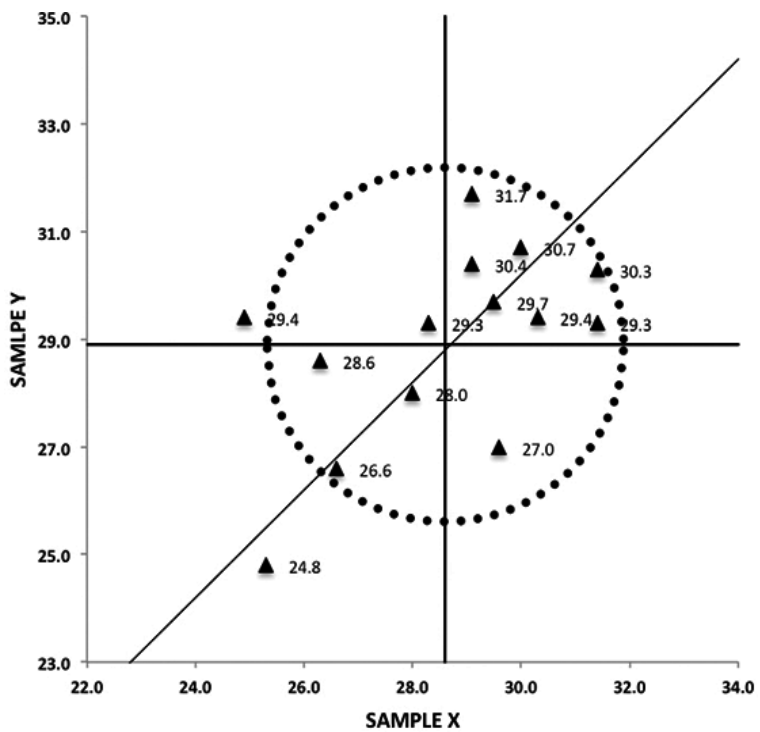

Figure 4. Youden plot (determination of PUFA, BF method).

If there were no systematic variations amongst laboratories, the pattern of points would be expected to be circular. The greater the systematic variations, the more elliptical the pattern will become.

The results obtained by the laboratories may now be compared by using the two methods as follows:

- Three laboratories (3, 6 and 9) are within the 95% circle and near each other on both plots.
- Five laboratories (4, 5, 8, 10 and 17) are within the 95% circle on both plots but widely separated from each other.
- Three laboratories (2, 14, 16) are outside the limit on one plot but not on the other.
- Three laboratories (11, 12 and 15) are outside the limit on both plots.
- Laboratory 13 is systematically low using both methods.
- The two standard deviations are approximately equal.

3.2. Determination of cholesterol levels in the blood

As part of a collaborative study to assess a new method that allows the determination of the total amount of cholesterol in the blood, two samples and the instructions to analyse each sample are sent to ten laboratories [36].

Table 3 shows the results obtained in mg total cholesterol per 100 mL of serum.

Figure 5 provides a two-sample plot of the results. The clustering of points suggests that the systematic errors of the analysts are significant. Two laboratories (1, 5) are outside the limit on the plot and widely separated from each other.

The vertical line at 248.9 mg/100 mL is the mean value for sample X, whereas the horizontal line at 246.5 mg/100 mL corresponds to the mean value of sample Y. To estimate σ_{rand} and σ_{syst}, the values for D_i and T_i are calculated first, followed by the standard deviations.

Because the F-ratio (2.530) is lower than F(0.05,9,9), which is 3.179, it is concluded that the systematic errors between the analysts are insignificant at the 95% confidence level.

If the true values for both samples are known, it is possible to test the presence of systematic errors. When there are no systematic method errors, the sum of the true values, μ_{tot}, for samples X and Y is equal to

$$\mu tot = \mu X + \mu Y \tag{16}$$

should fall within the confidence interval around *T*. A two-tailed *t*-test of the following null and alternate hypotheses is applied:

$$H0: T = \mu tot \; HA: T \neq \mu tot$$

http://dx.doi.org/10.5772/66411

Laboratory	Sample X	Sample Y
1	245.0	229.4
2	247.4	249.7
3	246.0	240.4
4	244.9	235.5
5	255.7	261.7
6	248.0	239.4
7	249.2	255.5
8	255.1	224.3
9	255.0	246.3
10	243.1	253.1

Table 3. Determination of cholesterol in serum (mg cholesterol/100 mL of serum).

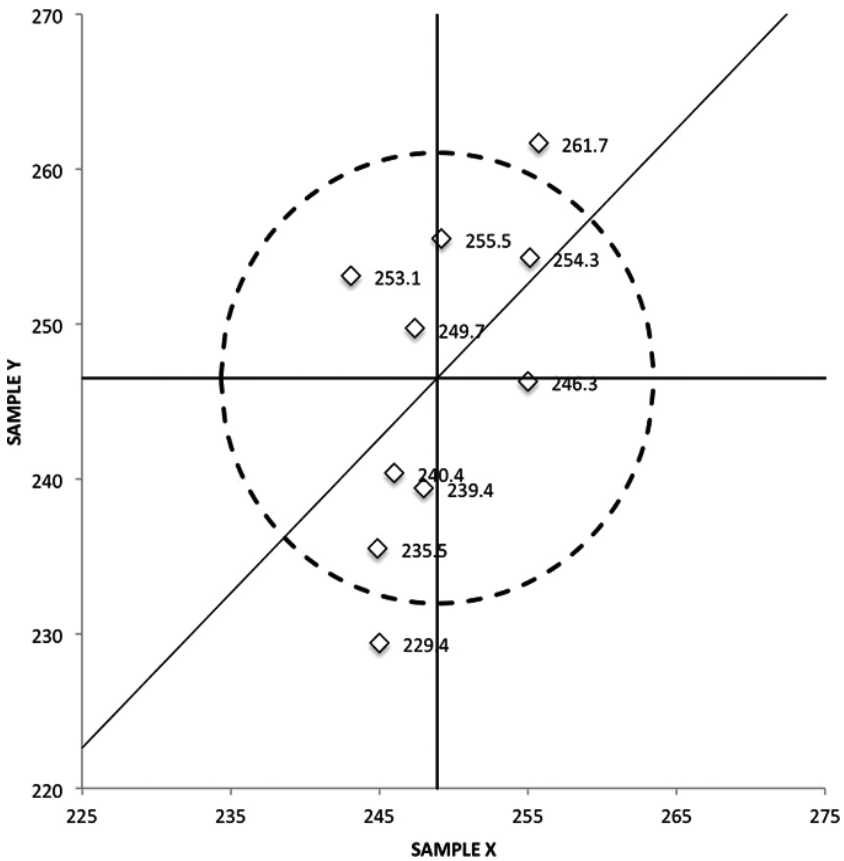

Figure 5. Youden plot (determination of cholesterol).

This occurs so as to determine if there is evidence for a systematic error in the method. The test statistic, *t*exp, is

$$t_{exp} = \frac{\left|\dot{T} - \mu_{tot}\right| \sqrt{n}}{S_T \sqrt{2}} \quad (17)$$

with $n-1$ degrees of freedom. The $\sqrt{2}$ in is included in the denominator because S_T underestimates the standard deviation when comparing T to μ_{tot}.

Because this value for *t*exp is smaller than the critical value of 2.26 for *t*(0.05, 9), there is no evidence for a systematic error in the method at the 95% confidence level.

3.3. Determination of aspirin in pharmaceutical preparations

The results of determinations, made in ten laboratories, on two similar aspirin preparations are given in **Table 4** [37].

Laboratory	Sample X	Sample Y
1	50.45	52.55
2	49.89	52.00
3	49.60	51.70
4	50.26	52.11
5	49.78	51.79
6	49.92	51.81
7	50.22	52.35
8	50.40	52.26
9	50.17	52.24
10	49.85	51.87

Table 4. Weight content (%) of aspirin in pharmaceutical preparations [37].

The analysis of data concerning materials X and Y reveals the following:

- A high level of interlaboratory variation.
- For each of the laboratories, the observed difference between the aspirin content of the two materials is approximately the same.

The average values for the materials (50.054% for X and 52.068% for Y) are used for the centroid. The results are plotted in **Figure 6** where it can be deduced that the data pointed on the diagram fall in either the first or the third quadrant (line X = Y). This is a consequence of the fact that laboratories which obtained "high" values for material X also obtained "high" values for material Y. The opposite is also true; laboratories reporting "low" values for X also reported "low" values for Y.

http://dx.doi.org/10.5772/66411

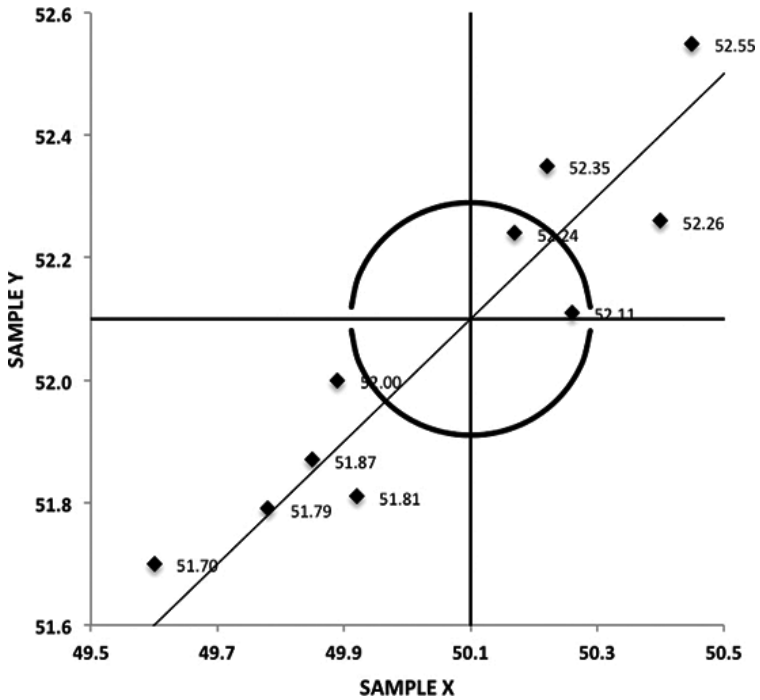

Figure 6. Youden plot (determination of aspirin).

Only two laboratories (2 and 9) are within the 95% circle. Because the F-ratio (25.834) is larger than F(0.05,9,9), which is 3.179, it is reasoned that the systematic errors between the analysts are significant at the 95% confidence level.

All this suggests that the variability in the difference between the two samples is evident. There are two factors which influence the variability of the differences. These are the random error of measurement and the heterogeneity of the materials. It is easy to infer that if the materials were relatively heterogeneous, then the reported differences would show considerable variability. That this is not so in this example is evidence that the laboratories were dealing with relatively homogeneous materials, i.e. the composition of the samples of materials X and Y received by each laboratory was essentially the same.

The larger the random error of measurement associated with an analytical procedure, the more varied the results of replicate measurements are. A relatively large random error of measurement will also cause the differences between them to vary considerably. The example presented herein indicates that the observed differences are approximately the same. This suggests that the random error experienced in each laboratory is relatively small. Perpendicular dispersions to the bisector, for homogeneous materials, are reflections of the within-laboratory variability. It is worth mentioning at this point that the example introduced here reveals that the within-laboratory variabilities are small, compared with the systematic between-laboratory variabilities.

Finally, it is sometimes believed that the centroid in a Youden diagram gives a good estimate of the true values of the two materials. It is the authors' view that this may often not

be the case. The scatter of the results obtained by several laboratories around the mean value is a result of both random within-laboratory and systematic between-laboratory variability. Averaging may result in the mean value being close to the true value. It may also result in a mean value which is, for example, much higher than the true value. Suppose, for example, that all the laboratories had a similar positive bias in one of the steps of the procedure. The result will be a scatter about a mean value which maybe higher than the true value.

3.4. Determination of methylparaben in surface water by liquid chromatography-negative electrospray ionization tandem mass spectrometry

Since parabens were discovered due to their antimicrobial activity, they have been widely used as bactericides, fungicides and preservative agents in many cosmetics, pharmaceuticals, personal care products and food, amongst other consumer products.

Although the toxicity of these compounds is very low, they present a weak estrogenic activity and are considered as endocrine disruptors. That is why they have been classified as emerging contaminants attracting scientific attention on a global scale. These compounds, after consumption, reach wastewater treatment plants, where they are not efficiently removed; thus, they end up in the environment. These chemical compounds consist of detergents, soaps and/or other products.

An analytical method for the determination of methylparaben (MeP) in surface water samples is applied during a period of 15 days.

The method is based on solid-phase extraction (SPE) and subsequent analysis by high-performance liquid chromatography-triple quadrupole mass spectrometry (HPLC-QqQ-MS) [38]. The Youden plot has been applied to the results. A detailed description of the completed experimental procedure is shown below:

3.4.1. Experimental part

3.4.1.1. Materials and reagents

HPLC-grade water, acetone and methanol were purchased by a company in Spain. Analytical-grade formic acid (98%), sulphuric acid (97%) and hydrochloric acid (37%) were acquired from another specialist industry in Spain. Ammonium acetate, MeP (≥99%), was bought from a firm in the USA.

Three millilitres SPE cartridges, packed with 60 mg of Oasis HLB, were purchased from Waters (Milford, MA, USA).

Stock solution, at a concentration of 1000 mg L^{-1}, was prepared in methanol and stored at 4°C. Working solutions were prepared by diluting the stock standard solutions in methanol.

3.4.1.2. Sample collection

Surface water samples were collected in May 2016 and were taken from Guadalquivir River (Seville, Spain). These samples were collected in amber glass bottles precleaned with acetone and methanol. In order to stabilize them, acetonitrile was immediately added after sampling

to achieve a final concentration of 0.5% v/v. Stabilized samples were stored at 4°C until further analysis, which was carried out within 48 h after sample collection. Prior to extraction, samples were filtered through a 1.2 μm glass-fibre membrane filter supplied by a British manufacturer.

3.4.1.3. Solid-phase extraction

Oasis HLB cartridges were conditioned using 3 mL of methanol followed by 3 mL of 0.5N hydrochloric acid and 3 mL of de-ionized water. Prior to extraction, the pH of the sample was adjusted to 2 by the addition of sulphuric acid 40% (v/v). The acidified sample (250 mL) was percolated through the cartridge at a flow rate of approximately 10 mL/min^{-1}. Then, the volumetric flask containing the sample was rinsed with 5 mL of de-ionized water, and the extract was added to the cartridge.

After loading the cartridges, they were washed with 5 mL of de-ionized water and dried for 10 min. The elution of the analytes was carried out with four successive aliquots of 1 mL of methanol at a flow rate of about 1 mL/min. The eluates were collected in 10-mL collection tubes and evaporated to dryness at room temperature by a gentle nitrogen stream. Finally, the extracts were reconstituted in 1 mL of methanol, filtered through a 0.45 μm nylon filter, and a 20-μL aliquot was injected into the HPLC instrument.

3.4.1.4. High-performance liquid chromatography-mass spectrometry

Separation was carried out using an Agilent 1200 series HPLC chromatography system equipped with a vacuum degasser, a binary pump, an autosampler and a thermostated column compartment. MeP was isolated with a Zorbax Eclipse XDB-C18 Rapid Resolution HT (4.6 mm × 50 mm i.d.; 1.8-μm particle size) column, using an isocratic elution with methanol (30%) and aqueous 5 mM ammonium acetate solution (70%) as mobile phase. Flow rate was 0.6 mL/min. The injection volume was 20 μL. The column temperature was maintained at 25°C.

The HPLC system was coupled to a 6410 triple quadrupole (QqQ) mass spectrometer (MS) equipped with an electrospray ionization source operating in negative mode. Two transitions were used for its identification (92.1m/z) and confirmation (136.1m/z).

The ionization of analytes was carried out using the following settings:

- MS capillary voltage 3000 V.
- Drying-gas flow rate 9 L/min^{-1}.
- Drying-gas temperature 350°C.
- Fragmentor 70 V.
- Collision energy 16 V.
- Nebulizer pressure 40 psi. Instrument control and data acquisition were carried out with MassHunter software.

3.4.2. Results and discussion

The results for the different days are shown in **Table 5**.

Because the F-ratio (3.000) is larger than F(0.05,14,14), which is 2.484, it is concluded that the systematic errors between the analysts are significant at the 95% confidence level.

Figure 7 depicts the results. By observing it, one may see that most of points fall in either the first or the third quadrant. Two days (5 and 11) are outside the 95% circle. The following comments may be drawn hereto:

– Days 4, 14 and 15 are within the 95% circle, near each other in the first quadrant.

– Days 2, 3 , 8, 10 and 13 are within the 95% circle, close together, in the third quadrant, although the latter is the very edge of the circle.

– Days 1, 9 and 12 are within the 95% circle, close together but in different quadrants.

– Day 7 is also within the circle but farther away from the other days.

– Day 6 is within the circle but at the very edge of it.

Finally, the z-score approximation is applied (pls. refer to Section 2.2) [39]. The results are shown in **Table 6**. A comparison of the results obtained by the laboratories using the Youden plot and Z-Score is done as follows:

Day	Area MeP (HPLC-MS/MS)		MeP concentration (ng/L)	
	X	Y	X	Y
1	3329	3574	16.1	17.8
2	3255	3388	15.6	16.5
3	3224	3302	15.3	15.9
4	3518	3886	17.4	20.1
5	3621	4095	18.2	21.6
6	3738	3435	19.0	16.8
7	3108	3862	14.5	19.9
8	3145	3200	14.8	15.2
9	3205	3531	15.2	17.5
10	3266	3133	15.6	14.7
11	3056	3076	14.1	14.3
12	3468	3537	17.1	17.6
13	3065	3162	14.2	14.9
14	3357	3758	16.3	19.1
15	3417	3877	16.7	20.0

Table 5. Determination of MeP in two similar surface water samples in different days.

http://dx.doi.org/10.5772/66411

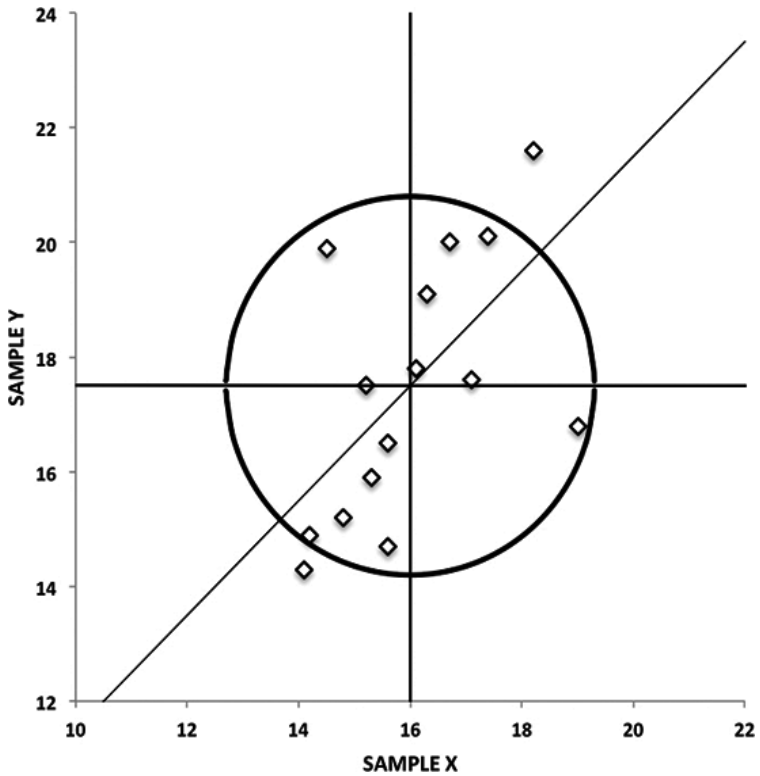

Figure 7. Youden plot (determination of MeP).

- The results obtained with the z-score are in line with those observed in the Youden plot.
- In both methods, days 5 and 11 whose systematic errors are relatively high compared to random errors showed unsatisfactory results.
- Day 6 is discarded using z-score but not on the Youden plot, although it is found very close to the boundary of the 95% circle.
- Days 1, 2, 3, 9, 12 and 14 showed satisfactory results with the z-score method applied.
- Days 4, 7, 8, 10, 13 and 15 showed questionable results. Day 7 is the furthest from the other days in the Youden Plot, and day 13 is on the edge of the 95% circle.

3.5. Antibody concentrations: an ISO-13528:2005(E) example

Finally, a confidence ellipse, calculated as described in ISO 13528 [40], has been used as an aid to interpret the plot so as to deal with those situations, in which the two samples differ in magnitude of the property measured. A Youden Plot for the original data may be derived from the z-scores (as explained below). It is constructed by plotting the z-scores obtained on one of the materials against the z-scores obtained on the other material.

Day	X	Y	ε_T	C_S	C_R	ε_S	ε_R	E_T	ε_R^2	Z-score	
1	16.1	17.8	0.54	0.53	0.11	0.45	0.09	0.54	0.01	0.52	Satisfactory
2	15.6	16.5	1.03	−0.78	0.67	−0.55	0.48	1.03	0.23	0.99	Satisfactory
3	15.3	15.9	1.67	−1.37	0.95	−0.99	0.68	1.67	0.46	1.60	Satisfactory
4	17.4	20.1	3.11	3.07	0.52	2.66	0.45	3.11	0.2	2.98	Questionable
5	18.2	21.6	4.77	4.65	1.06	3.88	0.89	4.77	0.78	4.57	Unsatisfactory
6	19.0	16.8	3.45	1.9	2.88	1.37	2.08	3.45	4.31	3.31	Unsatisfactory
7	14.5	19.9	2.62	0.87	2.47	0.68	1.94	2.62	3.76	2.52	Questionable
8	14.8	15.2	2.52	−2.29	1.07	−1.72	0.80	2.52	0.64	2.42	Questionable
9	15.2	17.5	0.44	−0.31	0.31	−0.22	0.22	0.44	0.05	0.42	Satisfactory
10	15.6	14.7	2.85	−2.02	2.02	−1.43	1.43	2.85	2.03	2.73	Questionable
11	14.1	14.3	3.59	−3.36	1.24	−2.62	0.97	3.59	0.93	3.44	Unsatisfactory
12	17.1	17.6	1.44	1.05	0.99	0.74	0.7	1.44	0.49	1.38	Satisfactory
13	14.2	14.9	3.01	−2.88	0.85	−2.32	0.69	3.01	0.47	2.88	Questionable
14	16.3	19.1	1.75	1.61	0.68	1.23	0.52	1.75	0.27	1.68	Satisfactory
15	16.7	20.0	2.`70	2.51	0.98	1.94	0.76	2.70	0.58	2.59	Questionable

Table 6. Application of z-score to the determination of MeP in two similar surface water samples.

For ease of reference, let A and B denote the two materials:

1. Calculate the averages and standard deviations of the two sets of data and the correlation coefficient ($\hat{\rho}$).
2. Calculate the z-scores for the two materials as follows:

$$Z_{A,i} = \frac{(X_{A,i} - \overline{X_A})}{S_A}; \quad and \quad Z_{B,i} = \frac{(X_{B,i} - \overline{X_B})}{S_B} \tag{18}$$

3. Calculate the combined scores for the two materials:

$$Z_{A,B,i} = \sqrt{Z_{A,i}^2 - 2\hat{\rho}\, Z_{A,i} Z_{B,i} + Z_{B,i}^2} \tag{19}$$

4. In terms of the standardized variables, the confidence ellipse may be written in terms of Hotelling's T^2:

$$Z_A^2 - 2\hat{\rho}\, Z_A Z_B + Z_B^2 = \left(1 - \hat{\rho}^2\right) T^2 \tag{20}$$

where

$$T^2 = 2\left\{\frac{(p-1)}{(p-2)}\right\} F_{(1-\alpha)}(2, p-1) \tag{21}$$

Here, $F_{(1-\alpha)}(2, p-1)$ is the tabulated (1-α)-fractile of the F-distribution with 2 and (p-1) degrees of freedom.

As recommended by the International Organization for Standardization (ISO), the ellipse may be drawn on a graph with the z-scores Z_A and Z_B as the axes by plotting a series of points for $-T < Z_A < T$ with

$$Z_B = \hat{\rho} Z_A \pm \sqrt{\left(1-\hat{\rho}^2\right)\left(T^2 - Z_A^2\right)} \tag{22}$$

To interpret the Youden Plot, the combined z-scores may be used. The highest combined z-score corresponds to the highest significance level of 100%.

Also, the combined z-scores aid to identify the outlying points.

When a Youden Plot is constructed, it may be interpreted as follows:

– If a point is well separated from the rest of the data, it means that the result is subject to bias because the laboratory did not follow the test method correctly. Points far away from the major axis could also represent laboratories showing a considerable variation and inadequate repeatability outcomes.

– A positive relationship between the results for the two materials indicates that there is a cause of between-laboratory variation that is common to many of the laboratories, suggesting that the methodology may not have been adequately specified. If the method is reproduced, it may lead to an overall improvement.

Table 7 shows data obtained by testing two similar samples for antibody concentrations and the calculations required to derive the confidence ellipse. With p = 29 laboratories and using

Row	Data		Z-score		Combined Z-score
	Allergen A (U)	Allergen B (U)	Z_A	Z_B	Z_{AB}
1	12.950	9.150	0.427	0.515	0.370
2	6.470	6.420	−1.540	−0.428	1.275
3	11.400	6.600	−0.043	−0.366	0.336
4	8.320	4.930	−0.978	−0.942	0.737
5	18.880	13.520	2.228	2.023	1.641
6	15.140	8.220	1.092	0.194	0.965
7	10.120	7.260	−0.432	−0.138	0.349
8	17.940	9.890	1.942	0.770	1.501
9	11.680	4.170	0.042	−1.204	1.234
10	12.440	7.390	0.272	−0.093	0.344
11	6.930	7.780	−1.400	0.042	1.430
12	9.570	5.800	−0.599	−0.642	0.477

Row	Data		Z-score		Combined Z-score
	Allergen A (U)	Allergen B (U)	Z_A	Z_B	Z_{AB}
13	11.730	5.770	0.057	−0.652	0.693
14	12.290	6.970	0.227	−0.238	0.429
15	10.950	6.230	−0.180	−0.493	0.388
16	10.950	5.900	−0.180	−0.607	0.497
17	11.170	7.740	−0.113	0.028	0.134
18	11.200	8.630	−0.104	0.335	0.415
19	7.640	3.740	−1.185	−1.353	0.985
20	12.170	7.330	0.190	−0.114	0.282
21	10.710	5.700	−0.253	−0.676	0.529
22	7.840	6.070	−1.124	−0.549	0.833
23	20.470	15.660	2.710	2.762	2.098
24	12.600	11.760	0.321	1.415	1.210
25	11.370	4.910	−0.052	−0.949	0.913
26	11.360	13.510	−0.055	2.019	2.059
27	10.750	5.480	−0.241	−0.752	0.607
28	12.210	9.770	0.203	0.729	0.603
29	7.490	5.820	−1.230	−0.635	0.902
Mean	11.543	7.659	0.000	0.000	
Standard deviation	3.294	2.897	1.000	1.000	

Units (U) in thousands (k) per litre (l) of sample, where a unit is defined by the concentration of an international reference material

Hotelling's T^2	6.927
T	2.632
$\hat{\rho}$	0.706
F (5%)	3.34

Table 7. Data and calculations on concentrations of antibodies for two similar allergens.

a significance level of 100% = 5%, $F_{(1-\alpha)}(2, p-1) = 3.34$. Hence, T = 2.632. The ellipse is shown, together with the points representing the z-scores, in **Figure 8**, in tandem with the ellipses pertaining to probability levels of 100% = 1% and 0.1%.

As depicted in **Figure 8**, laboratories 5 and 23, with combined z-scores of 1.641 and 2.099, respectively, are found in the top right-hand quadrant. Laboratory 26 has a high z-score on material B (2.019) compared to material A (−0.055) and a combined z-score of 2.059 followed by laboratory 8 with a combined z-score of 1.501. The points for laboratories 23 and 26 fall between the ellipses for the 5% and 1% probability levels. Thus, the results may be perceived as giving rise to warning signals.

http://dx.doi.org/10.5772/66411

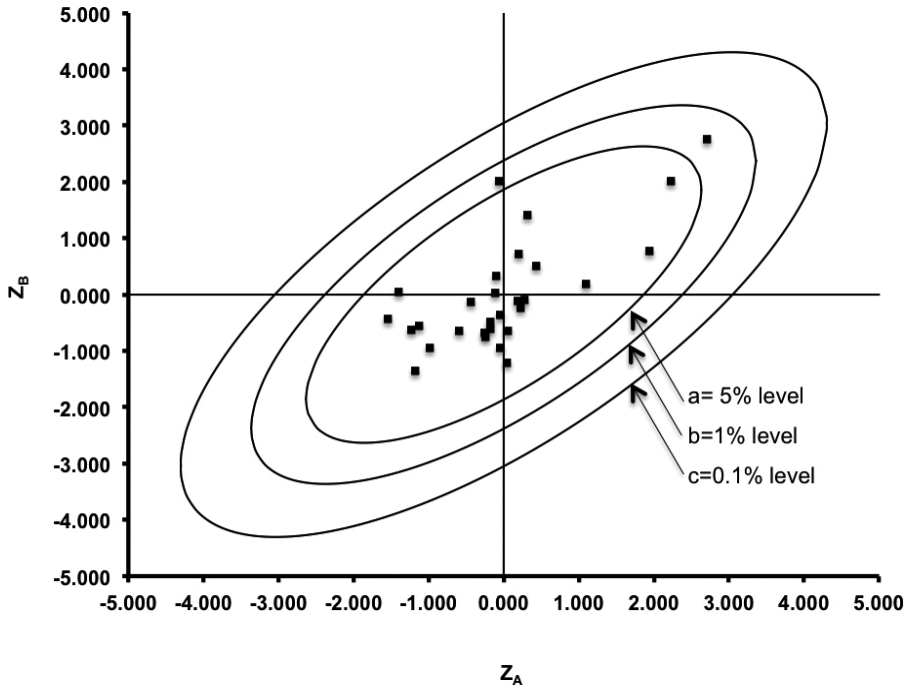

Figure 8. Youden plot of z-scores from **Table 7** (concentrations of antibodies for two similar allergens).

4. Conclusions

Intercomparison exercises are of great value in systems of quality assessment allowing the examination of the analytical process and the generated results. Youden plot or Youden analysis is particularly aimed at interlaboratory comparisons, obtaining accurate information without much effort. The main characteristic is its ability to separate the systematic and random errors with minimal effort on the part of the participants. To implement the method:

- Two similar materials (samples A and B) with small differences in the concentration of the characteristics (magnitude) are required with the purpose of determining their content [1993].
- A scatter plot is drawn in which the x-axis indicates one of the reported values and the y-axis the other, being the scale units the same along its axis. Each pair of results, corresponding to a given laboratory, is a point in the Youden plot 2.0.
- The points occur mainly in the (+ +) and (− −) quadrants, forming an elliptical pattern around a line bisecting these quadrants at a 45° angle, when systematic errors are larger than random errors. The circle centred at the intersection of the reported value medians (once outliers removed) affords a test on the randomness of results, its radius being a multiple of the within-laboratory standard deviation. The suitability and benefits of the Youden method have been applied to a number of results obtained from different fields such as food; clinical and pharmaceutical applications, in order to determine the concentration of

polyunsaturated fatty acids in fats and oils; total blood cholesterol; and aspirin in pharmaceutical preparations, respectively.

- In most cases, it is observed that systematic errors may be regarded as the main cause of variation with most of points in quadrants (+ +) and (− −). Finally, a detailed procedure for the determination of methylparaben in surface waters, of special relevance nowadays in the environmental field, has been developed by liquid chromatography-tandem mass spectrometry. In this experimental system, an alternative to the Youden method based on the z-score has also been assessed showing no discrepancy between both methods.
- A confidence ellipse is proposed in ISO 13528:2005 to deal with those situations where the two samples differ in magnitude of the property measured. The extension of the Youden method based on the confidence ellipse may be used as the building platform for further studies, incorporating amongst other three- and four-dimensional Youden Plots.

Abbreviation list

BF	Boron trifluoride-methanol
FDA	Food and Drug Administration
ISO	International Organization for Standardization
MeP	Methylparaben
MS	Mass spectrometer
PUFA	Polyunsaturated fatty acids
QqQ	Triple quadrupole
SPE	Solid-phase extraction

Author details

Julia Martín, Nieves Velázquez and Agustin G. Asuero*

*Address all correspondence to: asuero@us.es

Department of Analytical Chemistry, Faculty of Pharmacy, The University of Seville, Seville, Spain

References

[1] Funk W, Dammann V, Donnevert G. Quality assurance in analytical chemistry application in environmental, food and materials analysis, biotechnology and medical engineering. 2nd ed., Wiley-VCH: Weinheim, Germany, 2007; pp. 179–188.

[2] Horwitz W. Protocol for the design, conduct and interpretation of collaborative studies. Pure Appl. Chem. 1995; **67**(2):331–343.

[3] Horwitz W. Protocol for the design, conduct and interpretation of method-performance studies. Pure Appl. Chem. 1988; **60**(6):855–864.

[4] Taylor JK. Quality assurance of chemical measurements. CRC: Boca Raton, FL, p. 19.

[5] AMC Fitness for purpose: the key feature in analytical proficiency testing. Analytical Methods Committee, AMCTB No. 68, Anal. Methods. 2015; **7**:7404–7405.

[6] Kateman G, Buydens L. Quality control in analytical chemistry. Wiley: New York, 1993; 3.4.3 Youden Plot, pp. 132–136.

[7] Miller JN, Miller JC. Statics and chemometrics for analytical chemistry. Pearson: Harlow, England, 2010.

[8] González AG, Herrador MA, Asuero AG. Intralaboratory testing of method accuracy from recovery assays. Talanta. 1998; **48**:729–736.

[9] Ketterer ME. Assessment of overall accuracy of lead isotope ratios determined by inductively coupled plasma mass spectrometry using batch quality control and the Youden two- sample method. J. Anal. At. Spectrom.. 1992; **7**(7):1125–1129.

[10] Delwiche SR, Palmquist DE, Lynch JM. Collaborative studies for cereals analysis. Cereals Food World. 2005; **50**(1):9–17.

[11] Heyden YV, Smeyers-Verbeke J. Set-up and evaluation of interlaboratory studies. J. Chromatogr. A. 2007; **1158**(1–2):158–167.

[12] Hund E, Massart DL, Smeyers-Verbeke J. Interlaboratory studies in analytical chemistry. Anal. Chim. Acta. 2000; **423**(2):145–465.

[13] ISO 13528:2005 (E) Statistical methods for use in proficiency testing by interlaboratory studies. International Standardization Organization: Genève, 2005.

[14] AOCS Procedure M4-86. Evaluation and Design of Test Methods. Collaborative Study Procedures. The American Oil Chemists' Society; 2012. pp. 1–9.

[15] Boley N. Selection, Use and Interpretation of Proficiency Testing (PT) Schemes by laboratories – 2000. The Eurachem Nederland, Task Group "Proficiency Testing Schemes" and the Laboratory of the Government Chemist (LGC), United Kingdom.

[16] Pocklington WD. Guidelines for the development of standard methods of collaborative study: organization of interlaboratory studies and a simplified approach to the statistical analysis of collaborative study results, 5th ed.. Laboratory of the Government Chemist: Teddington, England, 1990.

[17] Wernimont GT, Spendley W. Use of statistics to develop and evaluate analytical methods. Assoc. Off. Anal. Chem: Arlington, VA, USA, 1987.

[18] Youden WJ, Steiner EH. Statistical manual of the association of official analytical chemists. Statistical techniques for collaborative tests. Planning and analysis of results of collaborative tests. AOAC: Washington DC, 1975.

[19] Juran JM, Godfrey AB. (Eds.) Juran's quality handbook, 5th ed.. McGraw-Hill: New York, 1999; Youden Two-Sample Plan, 47.28.–47.29.

[20] Zhong Z, Li G, Luo J, Chen W, Liu L, He P, Luo Z. Proficiency testing for determination of lead and arsenic in cosmetics: comparison of analytical procedures and evaluation of laboratory performances. Anal. Methods. 2015; **7**(7):3196–3177.

[21] Juneja A, Anand A.Understanding statistical concepts in laboratory quality control measures in biomedical research. Int. J. Res. Med. Sci.. 2015; **3**(11):3443–3445.

[22] Youden WJ. Experimentation and measurement. National Institute of Standards and Technology, NIST, Special Publication 672. U.S. Department of Commerce. Reprinted; 1977.

[23] Youden WJ. Graphical diagnosis of interlaboratory test results. Ind. Qual. Control. 1959; **15**(11):24–28.

[24] Youden WJ. Statistical aspects of the cement testing program. Proc. Am. Soc. Testing Mat. Philadelphia, PA.. 1959; **59**: 1120–1128.

[25] Youden WJ. Evaluation of chemical analyses on two rocks. Technometrics. 1959; **1**(4):409–417.

[26] Tracy ND, Young JC. A bivariate control chart for paired measurements. J. Qual. Technol.. 1995; **27**(4):370–376.

[27] Mandel J, Lashof TW. Interpretation and generalization of Youden's two- sample. J. Qual. Technol.. 1974; **6**(1):22–36.

[28] Cornell JA. The man and his methodology. AQQC Statistical Division Newsletter. 1993; **13**(2):9–18.

[29] Lavetz R. Statistical Manual. Chemical Proficiency Testing – NMI north ryde. Chemical & biological metrology. Document No 3. Australian Government. National Measurement Institute; 2014. pp. 1–15.

[30] Starink RJ and Visser RG. Interlaboratory studies: protocol for the organisation, statistics and evaluation. Institute for Interlaboratory Studies. Spijkenisse, The Netherlands; 2014. pp. 1–27.

[31] De Girolamo A, Ciasca B, Stroka J, Bratinova S, Visconti A, La anzio VMT. Determination of DON, FB1, FB2, ZEA, T-2, HT-2, OTA, AFB1, AFG1, AFB2, AFG2 in maize and determination of DON, ZEA, T-2, HT-2, OTA in wheat. Report of the 2014 Proficiency Test for LC-MS(MS) multi-mycotoxin methods. National Research Council of Italy. Institute of Sciences of Food Production. Italy; 2014.

[32] Kunsagi Z, Stroka J. Report on the validation of a method for the determination of Ochratoxin A in Capsicum spp. (Paprika and chilli). European Commission, Joint Research Centre, Institute for Reference Materials and Measurements: Geel, Belgium, 2012.

[33] Marchetto A, Mosello R, Tartari G. Atmospheric Deposition and Soil Solution Working Ring Test 2011. Laboratory ring test for deposition and soil solution sample analyses for the laboratories participating in the EU/Life + FutMon Project LIFE07 ENV/D/000218. Deliverable QA-Rwater-11. FutMon "Further Development and Implementation of an EU-level Forest Monitoring System". Consiglio Nazionale delle Ricerche. Istituto per lo Studio degli Ecosistemi. Verbania Pallanza; 2011. pp. 1–59.

[34] Villeneuve JP, de Mora SJ, Cattini C. World-wide and regional intercomparison for the determination of organochlorine compounds and petroleum hydrocarbons in mussel tissue IAEA-432. International Atomic Energy Agency, Marine Environment Laboratory: Monaco, Report No 74, March 2004.

[35] Sheppard AJ, Walting AE, Zmachinski H, Jones ST. Two lipoxidase methods for measuring cis, cis-methylene interrupted polyunsaturated fatty acids in fats and oils: collaborative study. J. Assoc. Offic. Anal. Chem.. 1978; **6**:1419–1423.

[36] Analytical chemistry 2.0.An electronic textbook for introductory courses in analytical chemistry. Chapter 15 quality assurance.Available from: http://acad.depauw.edu/harvey_web/eText%20Project/AnalyticalChemistry2.0.html.

[37] McCornic D, Roach A. Measurement, statistics and computation, analytical chemistry by open learning. ACOL, Wiley: Chichester, 1987; 4.7.4. Youden plot, pp. 322–325.

[38] Martín J, Camacho-Muñoz D, Santos JL, Aparicio I, Alonso E. Determination of emerging and priority industrial pollutants in surface water and wastewater by liquid chromatography–negative electrospray ionization tandem mass spectrometry. Anal. Bioanal. Chem.. 2014; **406**:3709–3716.

[39] Pérez A. Evaluación Estadística (Comparaciones Interlaboratorios - Análisis de Youden). Sociedad Española de Bioquímica Clínica y Patología Molecular. https://es.scribd.com/document/317778285/Estudios-Interlaboratorios-09-07-1.

[40] Jackson JE. Quality control methods for two related variables. Ind. Qual. Control. 1956; **7**: p. 2–6.

[41] Garrett RG. A comparison of Shewhart, Thompson and Howarth, and Youden plots – advantages and disadvantages. Explore Newsl. Assoc. Appl. Geochem.. 2015; **167**:5–11.

[42] Janik M, Yonehara H. The most recent international intercomparisons of radon and thoron monitors with the NIRS radon and thoron chambers. Radiat. Prot. Dosimetry. 2015; **164**(4):595–600.

[43] Zhou Q, Hu J, Li X, Li S, Gao Z, Xie W, Xu J. Comparison of traditional, trimmed traditional and robust Youden charts. Clin. Chim. Acta. 2015; **446**:213–217.

[44] Zhou Q, Hu J, Li X, Li S, Gao Z, Xu J, Xei W. Construction and application of the robust Youden plot in a EQA program. Accred. Qual. Assur. 2015; **20**(3):195–201.

[45] Van der Bresselaar AMHP, Abdoel CF, Ardanary D, van de Kamp G, Versluijs FAC. Preparation and control blood for external quality assessment of point-of-care international normalized ratio testing in the Netherland. Am. J. Clin. Pathol.. 2014; **14**:879–883.

[46] Xiao YL, Zhang CB, Zhao HJ, Kang FF, Wang W, Zhong K, Yuan S, Wang ZG. Application of ISO 13528 robust statistical methods for external quality assessment of blood glucose measurements in China. Accred. Qual. Assur.. 2014; **19**(5):397–401.

[47] O'Donnel GE, Hibbert DB. A study of the conditions of measurement required to evaluate bias in analytical results illustrated by the use of data from a multi-round, blind-duplicated, proficiency test. Analyst. 2013; **138**:3673–3678.

[48] Monteiro LR, Grafitti D, Albano F, Porfírio D, Fernandes Jr LP, Cotrim MEB, Pires MAF. Evaluation of a Brazilian ion chromatography interlaboratory study.Accred. Qual. Assur.. 2013; **18**(3):207–215.

[49] Shirono K, Iwase K, Okazaki H, Yamazawa M, Shikakume K, Fukumoto N, Murakami M, Yanagisawa M, Tsugoshi T. A study on the utilization of the Youden plot to evaluate proficiency test results. Accred. Qual. Assur.. 2013; **18**(3):161–174.

[50] Yu W, Tan Q, Yu L, Cao H. Comparison of statistical model of Youden plot for proficiency testing in ISO13528 and NATA. Metall. Anal.. 2013; **33**(12):74–80.

[51] Bremser W, Lücke FK, Urmetzer C, Fuchs E, Leist U. An approach to integrated data assessment in a proficiency test on the enumeration of *Escherichia coli*. J. Appl. Microbiol.. 2011; **110**(1):128–138.

[52] de la Calle Guntiñas MB, Semeraro A, Wysocka I, Cordeiro F, Quétel C, Emteborg H, Charoud-Got J, Linsinger TPJ. Proficiency test for the determination of heavy metals in mineral feed. The importance of correctly selecting the certified reference materials during method validation. Food Addit. Cont. Part A Chem. Anal. Control Expo. Risk Assess.. 2011; **28**(11):1534–1546.

[53] Kanefuji K, Tsugoshi T, Iwase K. Evaluation between statistical methods relate to the z score for use in proficiency testing. Bunseki Kagaku. 2011; **60**(7):571–577.

[54] Lavine BK. Learning Outcome Assessments in Quantitative Analysis Lab Using Youden Plots. New Trends in the Teaching of Analytical Chemistry. FACSS Analytical Science and Innovation. 2011 . https://www.scixconference.org/program/archive?p=4424&yearSelect=2011

[55] Mann I, Brookman B. Selection, use and interpretation of proficiency testing (PT) Schemes. 2nd ed. EA-Eurolab-Eurachem; 2011.

[56] Wang JS, Kee MK, Choi BS, Kim CW, Kim SS. Evaluation of external quality assessment results for HIV testing laboratories in Korea using current analytical methods. Clin. Chim. Acta. 2011; **412**(11–12):1127–1132.

[57] Heath E, Kosjek T, Farre M, Quintana JB, De Alecastro LF, Castiglioni S, Gans O, Langford K, Loos R, Radjenovic J, Rocca LM, Budzinski H, Tsipi D, Petrovic M, Barcelo

D. Second interlaboratory exercise on non- steroidal anti-inflammatory drug analysis in environmental aqueous samples. Talanta. 2010; **81**:1189–1196.

[58] Pankratov I, Elhanany S, Hening S, Zaritsky S, Ostapenlo I, Kuselman I. Development of a proficiency testing scheme for a limited number of participants in the field of natural water analysis. Accred. Qual. Assur.. 2010; **15**(8):459–466.

[59] de la Calle Guntiñas MB, Mysocka I, Quétel C, Vassileva E, Robouch P, Emteborg H, Taylor P. Proficiency test for heavy metals in feed and food in Europe. Trends Anal. Chem.. 2009; **28**(4):454–465

[60] Sanyal D, Rani A. Proficiency test for chemical laboratories for the analysis of a pesticide in a formulated product: interlaboratory study. J. AOAC Int.. 2009; **92**(1):271–278.

[61] Violante FGM, Bastos LHP, Cardoso MHWM, Rodrigues JM, Gourêa AV, Borges CN, Da Santos PR, Da Santos D, De A Goes HC, Souza VA, de Sâo J, Bandeira RDCC, Cunha V, Nóbrega A. Proficiency testing for the determination of pesticides in mango pulp: a view of the employed chromatographic techniques and the evaluation of laboratories performance. J. Chromatogr. Sci.. 2009; **47**(9):833–839.

[62] Wilke O, Horn W, Wiegner K, Jann O, Bremser W, Brödner D, Kalus S, Juritsch R, Till C. Investigations for the improvement of the measurement of volatile organic compounds from floor coverings within the health-related evaluation of construction products. BAM, Research-number ZP 52-5-20.49.1-1251/07. By Fraunhofer IRB Verlag, ISBN: 3-8167-8253-1, 2009.

[63] Bellis DJ, Hetter KM, Verostek MF, Parsons PJ. Characterization of candidate reference materials for bone lead via interlaboratory study and double isotope dilution mass spectrometry. J. Anal. At. Spectrom.. 2008; **23**(3):289–308.

[64] Flores L, Santo C, Trías M. Interlaboratory system to ensure and improve the quality of glassware calibration and use in a large laboratory. J. AOAC Int.. 2008; **91**(1):247–251.

[65] Gavlick WK, Tomkins DF. An updated liquid chromatographic assay for the determination of glyphosate in technical material and formulations. J. AOAC Int.. 2008; **91**(1):1–4.

[66] Nelsen TC, Wehling P. Collaborative studies for quantitative chemical analytical methods. Cereals Food World. 2008; **53**(5):285–288.

[67] Christopher SJ, Pugh RS, Ellisor MB, Mackey EA, Spatz RO, Porter BJ, Bealer KJ, Kucklick JR, Rowles TK, Becker PR. Description and results of the NIST/NOA a 2005 interlaboratory comparison exercise for trace elements in marine mammals. Accred. Qual. Assur.. 2007; **12**(3):175–187.

[68] Cooper TG, Hellenkemper B, Nieschlag E. External quality control for Semen analysis in Germany. J. Reproduct. Med. Endocrin.. 2007; **4**(6):331–335.

[69] Hare LB. It's not always what you say, but how you say it. Qual. Progress.. 2007; **40**(8):64–66.

[70] Lau A. Practical advice on Youden plot. Qual. Progress. 2007; **40**(10):10.

[71] Miller EL, Bimbo AP, Barlow SM, Sheridau B. Repeatability and reproducibility of determination of the nitrogen content of fishmeal by the combustion (Dumas) method and comparison with the Kjeldahl method: interlaboratory study. J. AOAC Int.. 2007; **90**(1):6–20.

[72] Siekmann L, Breuer H. Determination of cortisol in human plasma by isotope dilution-mass spectrometry. Definitive methods in clinical chemistry, I. J. Clin. Chem. Clin. Biochem.. 1982; **20**(12):883–892.

[73] Sorensen K, Grung M, Röttgers R. An intercomparison of in vitro chlorophyll A determinations for MERIS level 2 data validation. Int. J. Remote Sensing-MERIS. 2007; **28**(3–4):537–544.

[74] Bons, JA, de Boer, D, van Dieijen-Visser, MP and Wodzig, WK. Standardization of calibration and quality control surface-enhanced laser desorption/ionization time of flight mass spectrometry. Clin. Chim. Acta. 2006; **366**:249–256.

[75] Schilling P, Powilleit M, Uhlig S. Chlorophyll-a determination: results of an interlaboratory comparison. Accred. Qual. Assur. 2006; **11**(8):462–469.

[76] Svegl F, Strupi JS, Svegl IG. Proficiency testing of chloride content in different types of Portland cement. Accred. Qual. Assur. 2006; **11**(8):414–421.

[77] Costantini S, Ciaralli L, Ciprotti M, D'Ilio S, Giordano R, Mosca M, Sepe A, Jenofonte O. The network of the Italian laboratories: a proficiency test on the quantification of trace elements in serum. Ann. Ist. Super. Sanita. 2005; **41**(1):171–179.

[78] González AG, Herrador MA, Asuero AG. Practical digest for evaluating the uncertainty of analytical assays from validation data according to the LGC/VAM protocol. Talanta. 2005; **65**(4):1022–1030.

[79] Jones FE. Youden analysis of Karl Fisher titration data from an interlaboratory study determining water in animal feed, grain and forage. J. AOAC Int.. 2005; **88**(6):1840–1841.

[80] Siekmann L. Establishing measurement traceability in clinical chemistry. Accred. Qual. Assur.. 2004; **9**(1):5–17.

[81] Margolis SA, Vangel M, Duewer DL. Certification of standard reference material 970, ascorbic acid in serum, and analysis of associated interlaboratory bias in the measurement process. Clin. Chem.. 2003; **49**(3):463–469.

[82] Sorensen K, Grung M, Röttgers R. An intercomparison of in vitro chlorophyll A determination – preliminary results, Proceedings of the ENVISAT Validation Workshop (ESA SP-531), 9–13 December, Frascati, Lacoste, H. (Ed.), Italy, 2003.

[83] Official Methods of Analysis (AOAC). Guidelines for collaborative study procedures to validate characteristics of a method of analysis; 2002 AOAC International; Appendix D, pp. 2–12.

[84] Gadmar TC, Vogt RD, Osterhus B. The merits of the high-temperature combustion method for determining the amount of natural organic carbon in surface freshwater samples. Inter. J. Environ. Anal. Chem. 2002; **82**(7):451–461.

[85] Grijsen JG. Findings of third inter-laboratory AQC exercise. Inter-Laboratory AQC Exercise. Technical Assistant. Hydrology Project. Government of India & Government of The Netherlands; 2002. pp. 1–24.

[86] Berman S. Seventh intercomparison exercise on trace metals in sea water. Institute for National Measurement Standards National Research Council. Marine Chemistry Working Group. International Council for the Exploration of the Sea. Denmark. ICES Cooperative Research Report No. 237; ISSN 1017-6195. 2000.

[87] Carvhalo FP, Villeneuve JP, Cattini C. The determination of organochlorine compounds and petroleum hydrocarbons in a seaweed sample: results of a world-wide intercomparison exercise. Trends Anal. Chem.. 1999; **18**(11):656–664.

[88] McClure FD. A statistical evaluation of the Youden matched-pairs procedure. J. AOAC Int.. 1999; **82**(2):375–381.

[89] ISO 5725-5:1998. Accuracy (trueness and precision) of measurement methods and results – Part 5: alternative methods for the determination of the precision of a standard measurement method, 1998. https://www.iso.org/

[90] Marchetto A, Bianchi M, Geiss H, Muntan H, Serrini G, Serrini-Lanza G, Tartari GA, Mosello R. Performances of analytical methods for freshwater analysis assessed through intercomparison exercises. Mem. Ist. Ital. Idrobiol. 1997; **56**:1–13.

[91] Petersen PH, Ricós C, Stöckl D, Libeer JC, Baadenhuijsen H, Fraser C, Thienpont L. Proposed guidelines for the internal quality control of analytical results in the medical laboratory. Eur. J. Clin. Chem. Clin. Biochem. 1996; **34**(12):983–999.

[92] Wu GB, Meng H. Application and improvement of the Youden analysis in the intercomparison between flowmeter calibration facilities. Flow Measurement Inst. 1996; **7**(1):19–24.

[93] Feinberg M. Basic of interlaboratory studies: the trends in the new ISO 5725 standard edition. Trends Anal. Chem. 1995; **14**(9):450–457.

[94] Hewitt AD, Grant CL. Round-Robin study of performance evaluation soils vapor-fortified with volatile organic compounds. Environ. Sci. Technol. 1995; **29**:769–774.

[95] Mosello R, Bianchi M, Geiss H, Marche o A, Serrini G, Serrini-Lanza G, Tartari GA, Muntau H. AQUACON-MedBas Subproject No. 6. Acid rain analysis. Intercomparison 1/94. Join Research Centre European Commission, Rep. EUR 163332 EN, Istituto Italiano di Idrobiologia, National Research Council, Italy; 1995. p. 48.

[96] Barnes D, Dent G. Polystyrene films as a performance check for FTIR spectrometers. Spectrosc. Eur. 1994; **6**(2):8–14.

[97] Gaskin JE. Graphical diagnosis of interlaboratory quality control data for surface water samples. Analyst. 1994; **119**:1531–1535.

[98] Horwitz W. Nomenclature of interlaboratory analytical studies. Pure Appl. Chem. 1994; **66**(9):1903–1911.

[99] ISO 5725-2:1994. Accuracy (trueness and precision) of measurement methods and results – Part 2: basic method for the determination of repeatability and reproducibility of a standard measurement method. 1994. https://www.iso.org/

[100] Stephens RD, Rappe C, Hayward DG, Nygren M, Startin J, EsbØll A, Carlé J, Yrjänheikki EJ. World health organization international intercalibration study on dioxins and furans in human milk and blood. Anal. Chem.. 1992; **64**(24):3109–3117.

[101] Mesley RJ, Pocklington WD, Walker RF. Analytical quality assurance. A review. Analyst. 1991; **116**(10):975–990.

[102] Jones NE. Multiway analysis of variance for the interpretation of interlaboratory studies. Anal. Chem.. 1990; **62**:1532–1536.

[103] Oosting E, Neugebauer E, Keyzer JJ, Lorenz W. Determination of histamine in human plasma: the European external quality control study 1988. Clin. Exp. Allergy. 1990; **20**:349–357.

[104] Youden WJ. Classic paper. The collaborative test. (Reprinted from J-Assoc-Off-Agric-Chem, Vol. 46, PG 55-62, 1963). J. Assoc. Off. Anal. Chem.. 1990; **73**(2):194–201.

[105] Thompson M. Robust statistic and functional relationship estimation for comparing the bias of analytical procedures over extended concentration ranges. Anal. Chem.. 1989; **61**(17):1942–1945.

[106] Ferrus R, Torrades F. Bias-free adjustment of analytical methods to laboratory samples in routine analytical procedures. Anal. Chem.. 1988; **69**(13):1281–1285.

[107] Abern AM, Garrell RL. Exchange of comments on a new technique in chemical assay calculations. Anal. Chem.. 1987; **59**(23):2816–2818.

[108] Bauer CF, Grant CL, Jenkins TF. Interlaboratory evaluation of high-performance liquid chromatographic. Determination of nitroorganics in munition plant wastewater. Anal. Chem.. 1986; **58**(1):176–182.

[109] Boyer KW, Horwitz W, Albert R. Interlaboratory variability in trace element analysis. Anal. Chem.. 1985; **57**(2):454–459.

[110] Currie LA. The limitations of models and measurements as revealed through chemometric intercomparison. J. Res. Natl. Bur. Stand. 1985; **90**(6):409–419.

[111] Cleveland WS and McGrill R.The many faces of a scatterplot. J. Am. Stat. Assoc.. 1984; **79**(388):807–822.

[112] Jenkins TF, Leggett DC, Grant CL, Bauer CF. Reversed-phase high-performance liquid chromatographic determination of nitroorganics in munitions wastewater. Anal. Chem.. 1986;**58**:170–175.

[113] Mcdonald RW and Nelson H. A laboratory performance check for the determination of metals (Hg, Zn, Cd, Cu,Pb) in reference marine sediments. Canadian Technical Report of Hydrography and Ocean Sciences, No. 33. Institute of Ocean Sciences Department of Fisheries and Oceans, Sidney; 1984.

[114] Smith R. Organization and evaluation of interlaboratory comparison studies among southern African water analysis laboratories. Talanta. 1984; **31**(7):537–545.

[115] Jeffcoate SL. Use of Youden plot for internal quality control in the immunoassay laboratory.Ann. Clin. Biochem. 1982; **19**(6):435–437.

[116] Usui T. Quality control data evaluation: tolerance ellips expression of Youden plot. Japan. J. Clin. Chem.. 1982; **11**(2):98–103.

[117] Berman GA. Testing laboratory performance: evaluation and accreditation. NBS Publication 591, National Bureau of Standards: Gaithersburg, MD, U.S.A 1980.

[118] Evans WH. Qualification of estimates for total trace elements in food stuffs using measurement by atomic- absorption spectrophotometry. Analyst. 1978; **103**:452–468.

[119] Green A and Naegele R. Development of a system for conducting inter-laboratory tests for water quality and effluent measurements. EPA-600/4-77-031, Environmental Monitoring and Support Laboratory. Office of Research and Development. U.S. Environmental Protection Agent. Cincinnati, Ohio; 1977.

[120] Egan H. Collaborative analysis and the standardization of analytical methods. Proc. Society Anal. Chem.. 1972; **9**(11):245–249.

[121] Youden WJ. Graphical diagnosis of interlaboratory test results. (reprinted from Industrial Quality Control, XV, No 11, May 1959). J. Qual. Technol. 1972; **4**(1):29–33.

[122] Shendzel LP, Youden WJ. Systematic versus random error laboratory surveys. Am. J. Clin. Pathol.. 1970; **54**(3):448–450.

[123] Ku HH. Precision measurement and calibration. Statistical concepts and procedure. NBS Special Publication 300 –Volume 1, National Bureau of Standards: Washington, DC, 1969.

[124] Shendzel LP, Youden WJ. A graphic display of interlaboratory test results. Am. J. Clin. Pathol.. 1969; **51**(2):161–165.

[125] Schulz G. Determination of systematic and accidental errors of analytical procedure by Youden method. Zellstoff Papier. 1967; **16**(9):281.

[126] Youden WJ. Collaborative test. J. Assoc. Off. Agric. Chem.. 1963; **46**(1):55–62.

[127] Youden WJ. The sample, the procedure and the laboratory. Anal. Chem.. 1960; **32**(13):23A–27A.

[128] Linning FJ, Mandel J, Peterson JM. A plan for studying the accuracy and precision of an analytical procedure. Anal. Chem.. 1954; **26**(7):1102–1110.

[129] Wernimont GT. Design and interpretation of interlaboratory studies of test methods. Anal. Chem.. 1951; **23**(11):1572–1576.

Chapter 4

Using Lot Quality Assurance Sampling to Monitor the Prevalence of Abortions and the Quality of Reproductive Health Care in Armenia

Joseph J. Valadez and Lusine Mirzoyan

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66092

Abstract

Monitoring abortion prevalence is essential to plan control efforts. Lot Quality Assurance Sampling (LQAS) is an inexpensive, reliable method for monitoring abortion prevalence and access to quality reproductive health (RH) services. This chapter presents survey results from 2000 in three sites of Armenia (Gyumri, Gavar and Goris) using LQAS principles (i.e., 44%, 95% CI: ±6% of women had an induced abortion in their lifetime, a total abortion rate (TAR) of 2.0 abortions per woman). Modern contraceptive use was lowest in Goris (16%. 95% CI: ±7%) and highest in Gyumri (43%, 95% CI: ±11%). Only 37% (95% CI: ±9%) of women with an induced abortion received family planning information and 21% (95% CI: ±4%) of mothers were counselled about family planning after delivery. While limited access to family planning information and contraceptives is still an issue in Armenia, recently new reproductive health priorities—such as infertility, sex-selective abortions and abortions due to socio-economic difficulties—have become more common and can be investigated using LQAS in both community surveys and health facility assessments. This study demonstrates that measuring national abortion prevalence and access to services mask underlying variations; the awareness of which is essential for health program planning.

Keywords: abortions, reproductive health, Armenia, lot quality assurance sampling (LQAS), surveys

1. Introduction

According to the World Health Organization (WHO), an estimated 225 million women in developing countries wish to postpone or prevent bearing a child, but do not use any modern means of contraception [1]. Having become pregnant, many women terminate their pregnancy. In developing countries, 56% of all abortions are unsafe, compared with just 6% in the developed world [2]. Although legal and illegal medication abortions have become more common and have likely contributed to reductions of severe morbidity and maternal death, unsafe abortions in low- and middle-income countries still remain among the leading causes of maternal morbidity and mortality. The ethical aspects of terminating a pregnancy for non-medical reasons [3–5] are surely another concern.

Abortions, especially when repeated, pose many health risks to a woman, including the risk of death. Women who choose an abortion are approximately four times more likely to die in the following year than women who carry their pregnancies to term [6]. Studies of records for an entire female population in Denmark showed that the risk of death remains higher in each of the first ten years following the abortion. Moreover, the risk of death increases with each abortion: 45% after one abortion, 114% after two abortions and 192% after three or more abortions [7]. The risk of cervical cancer is 2.3 times higher in women with a history of one abortion and 4.92 times higher in women with a history of two or more abortions as compared to women with no history of abortion. Other studies have found a similar trend for the risks of subsequent ovarian and liver cancers; the elevated risk exists for both single and multiple abortions [8–11]. Approximately, 10% of women who abort will suffer immediate complications, of which approximately one-fifth are considered life-threatening [12]. The most common complications after abortions in later pregnancies include placenta previa, pre-term deliveries, handicapped newborns and ectopic pregnancies [13].

Abortions in the former Soviet states traditionally have been the main method of birth spacing, and Armenia, the focus of this chapter, has a similar profile in this regard [14]. According to the Armenia Demographic and Health Survey (ADHS) 2010, despite the decline over 10 years in the proportion of pregnancies resulting in induced abortion (a reduction from 55% in 2000 to 45% in 2005, and to 29% in 2010), 31% of all women of reproductive age reported having at least one induced abortion, with approximately two-thirds having had more than one abortion in their lifetime [15, 16]. The total abortion rate (TARe[1]) for Armenia in 2010 was 0.8 abortions per woman, which is lower than 1.8 in 2005 and significantly lower than 2.6 in 2000; this trend, however, might be attributed to the significant decline in the pregnancy rate over the same period [16]. This report highlights the fact that with the introduction of ultrasound determination of an unborn child's sex, gender-selective abortions became a common procedure in Armenia. In ADHS 2010, 8% of women having an abortion in the three years prior to the survey reported the desire to deliver a boy as the main reason for the termination.

[1] Total abortion rate (TAR) is "the total number of abortions a woman will have in her lifetime if current levels persist. This lifetime risk is a cohort measure and can be calculated with age-specific abortion rates or approximated by multiplying the abortion rate by the length of the reproductive period (30–35 year range) (MEASURE Evaluation: http://www.cpc.unc.edu/measure/prh/rh_indicators/specific/pac/abortions-per-1-000-live-births).

http://dx.doi.org/10.5772/66092

UNFP's report concurs by documenting that "with 114 boys born for every 100 girls in 2012, Armenia has one of the highest sex imbalance levels in the world…The imbalance is particularly dramatic for third births: the record level of 173 sons born for every 100 daughters has no known equivalence anywhere else in the world" [17]. Gender-specific abortion is particularly a concern in countries like Armenia having a low total fertility rate (TFRe[2]). Being below replacement levels at 1.7 live births per woman in 2010 the gender-imbalance can result in an increased reduction in the population size [18, 19].

As the TAR has important policy and demographic implications, low- and middle-income countries require an inexpensive and reliable method to monitor TAR, as well as the means to monitor the use of modern family planning methods and the availability and quality of reproductive health (RH) services. However, when measured at a national level, TAR can mask the sub-national variations that exist. Such variations affect policy formation and should be identified to assure the needs of differing subpopulations are addressed [20]. Lot Quality Assurance Sampling (LQAS) is one method that can be used locally for routine monitoring that produces reliable data for managing health programs and for improving health services [21]. With respect to preventing abortions in Armenia, LQAS can serve policy makers and planners by providing much-needed information specific for each region of the country [22, 23].

Although LQAS surveys can produce data of comparable quality to the demographic and health surveys (DHS) [24], they are intended to complement those surveys rather than compete with them. While DHS and multi-country indicator surveys (MICS) are international gold-standard survey methods, they are costly and they measure national indicators at approximately five-year intervals. LQAS is used for a related but different purpose: to frequently track priority indicators related to program coverage, the use of services and/or the quality of care in order to improve program services. Although LQAS data can be aggregated to produce prevalence measures for indicators used at the central level for policy making, its major advantages are to detect the variations in coverage at the sub-national level and to determine if targets and standards of care are met [25, 26].

While LQAS surveys are routinely implemented in many countries of Latin America, Africa and East and Southern Asia to assess a variety of public health programs and the quality of health care, they are less often used in the countries of Eastern Europe, South Caucasus or Central Asia [27, 28]. This chapter reports on a reproductive health survey carried out in three sites in three different regions of Armenia. While this survey was carried out in 2000, it is one of the first documented studies of abortion prevalence in the South Caucasus, and its results are still timely and pertinent to current global health issues.

The study was conducted as a baseline assessment for the USAID-funded Network for Health in Armenia project, which aimed to increase women's access to reproductive health and healthy family information and services through a coordinated effort of three private voluntary

[2] Total fertility rate (TFR) is the number of children the average woman would bear in her lifetime if she experienced the currently observed age-specific fertility rates throughout her reproductive years. The fertility rates refer to the three-year period before the survey (Armenia DHS 2010).

organizations (PVOs), government agencies and non-governmental organizations (NGOs). CARE International in Gyumri, Save the Children in Gavar and the Adventists Development and Relief Agency (ADRA) in Goris directed the three pilot sites (or program catchment areas). The pilot sites formed the nucleus of the regional networks and were linked to the central network in the capital under the leadership of ADRA.

2. Study methodology and results

2.1. Sampling method

The survey design used the LQAS method, which is an established analysis technique, originally developed as a classification method for industrial quality control during the 1920s and adapted to health sciences in the mid-1980s [29, 30] to manage units in a health system charged with delivering services, referred to in LQAS as supervision areas (SAs). In each SA, a random sample ('n') of individuals is assessed. A decision rule or cut-off value, designated as 'd', is selected to optimize identification of low performance SAs for a specified indicator. The 'd' depends on the sample size, thresholds for classifying high and low performance and the choice of two misclassification errors: the risk of overestimation when low coverage exists (β error) and the risk of underestimation when high coverage exists (α error). The upper threshold, 'p-Upper or pU', for identifying acceptably performing SAs is often suggested by program stakeholders, and the low threshold, 'p-Lower or pL', is usually 30% lower.

In the surveys used for this chapter, 19 locations were selected in each Network SA. The SA sample size, n = 19, guarantees that both misclassification errors do not exceed 0.10, as all corresponding values of 'pL' are 30% below 'pU'. SA having pL< p <Pu are classified according to their proximity to either pL or pU. SA with an intermediate performance has an equal chance of being classified as either high or low. The strength of LQAS is that it is very effective for identifying areas that are among the *worst of the worst* in terms of the indicator of interest. As an example, let us assume that the target for an indicator (such as exclusive breastfeeding) is set at 80%. The upper threshold, 'pU', is 80% for identifying high (or acceptably) performing SAs, while the lower threshold, 'pL', is 50%. The sample size 'n=19' and decision rule 'd=13' are selected to ensure α-errors of ≤10%, and β-errors ≤ 10%—or more formally:

- $P (X<d \mid n, pU \geq 80\%) \leq \alpha \leq 0.10$
- $P (X\geq d \mid n, pL \leq 50\%) \leq \beta \leq 0.10$
- $\alpha + \beta < 0.20$

The decision rules and error terms associated with varied sample sizes are found in [23, 24].

In Armenia, our intention was to apply LQAS for recurrent community monitoring by each of the three Network's members and the Network as a whole. The program area in each of the three sites was divided into administratively meaningful units or supervision areas. In each case, the Network member SAs were city areas, towns and surrounding villages. A supervisor —such as a doctor, or a nurse, or a midwife—managed each SA. With LQAS, data collectors

http://dx.doi.org/10.5772/66092

randomly select small samples (n=19) in each SA, which they use to judge performance. These data, when aggregated for a PVO or for the Network, formed a stratified random sample of the members' catchment area. All catchment areas in aggregate form the Network catchment area.

SAs with intermediate performance are classified as high or low depending on how close they fall to the relevant thresholds. There are three major advantages of using LQAS as compared to other probability sampling methods.

First, in addition to permitting calculation of a conventional average coverage for a program area, the method allows program managers to determine the relative performance of the different SAs that comprise the catchment area. For example, a typical health program area could include several communities with a total population of several thousand people. In baseline surveys, the supervisors determine whether any SA is below average and, therefore, needs special assistance. Based on baseline results, the annual (or semi-annual) performance benchmarks are established for recurrent monitoring. Then, in monitoring, the LQAS is used to determine whether SAs reach these performance benchmarks.

Second, LQAS uses a small sample size for making judgements. In addition to necessitating fewer interviews than for other conventional sampling methods, the smaller sample size leads to a quicker analysis and interpretation by local managers. For most applications, a sample of 19 individuals is required in each SA to judge whether it is below average or has reached a service delivery target. For calculation of a coverage proportion for the catchment area, the individual SA level results produced with the samples of 19 are aggregated while being weighted by the population size of the SA. Assuming there are five SAs, the total sample size would be 95. With p=50%, this sample results in a coverage measure for the catchment area having a confidence interval that does not exceed ±10% of the true coverage. Ideally it should be five or more SAs; however, having four SAs is also acceptable since the confidence interval increases only slightly (≤11%).

Third, as LQAS uses a small sample to judge whether a health worker's performance reaches a predetermined standard, data collection in densely populated areas does not seriously compete for time health workers can allocate to other health care activities. Hence, it is a practical management tool that is ideal for recurrent data collection.

2.2. Study procedures

A multi-stage sampling procedure was used to select the respondents. The first stage was a probability proportional to size (PPS) sampling method, meaning the larger the relative size of the community, the higher its chance of being selected. To carry out a PPS, a sampling frame was constructed, consisting of a list of all the communities in each SA with their population size. Each community population size is added to the next one to create a series of cumulative population sizes. So, if the first community has 453 people and the second one has 500, the first cumulative population size is 953 (or 453+500). Once all community population sizes have been summed, the sampling interval is calculated by dividing the total by the sample size (n=19 in this case). Assuming that the total population size is 43,518; when divided by 19 the result

is 2290.42, which is the sampling interval. A random number between 1 and 2290 is selected using a random number table to identify the first of 19 random interview locations; for argument's sake, let us say this number is 1406. Subsequently, the sampling interval is added to the 1406 to identify the second interview location, which results in 1406 + 2290.42 = 3696.42.

The third location is 3696.42 + 2290.42 = 5986.84. This process is repeated until all 19 interview locations are identified. One then looks at the cumulative population size in the sampling frame and identifies the community in which the 1406th person resides and then where the 3696th person resides and so forth. This is done for each of the 19 numbers. An Excel spread sheet is often used to automate selection of the communities using PPS.

The second stage is a random selection of the starting household using segmentation sampling [31]. This was done either by using a government map of a community or a city area or by drawing the map of a community while in the field at the time of interviewing. The idea was to avoid having to draw a map of the entire community showing the location of each household. Instead, the community was divided into segments of approximately equal sizes using the landmarks.

The essential point is that each segment should have approximately the same number of households. All segments were then numbered, and one of them was randomly selected. If the selected segment was of a manageable size having 30 or fewer households, then the interviewers drew a detailed map of this segment depicting all households. If the selected segment was still large, it was further divided into sub-segments to reach a manageable population size. In the next step, the households were enumerated and one was randomly selected.

In the third stage, a respondent is selected in the household. The respondents in this study, which are referred to herein as our client group, belonged to populations associated with the content of the three questionnaires and linked with the main health services under assessment. Each client group was sampled independently in each SA and had n=19 interviews:

- Mothers with children 0–11 months, interviewed to assess:
  - Antenatal care and safe delivery
  - Breastfeeding and child nutrition
  - Post-delivery family planning
- Women of reproductive age 15–49 years and not pregnant, interviewed to assess:
  - Use of modern contraceptives
  - Abortion-seeking behaviour
  - Post-abortion information provision
  - Knowledge of danger signs during pregnancy, delivery, after delivery
  - Knowledge of danger signs in a newborn
  - Knowledge of HIV and preventative behaviour

http://dx.doi.org/10.5772/66092

- Men of reproductive age 15–54 years, interviewed to assess:
  - Knowledge of danger signs during pregnancy, delivery, after delivery
  - Knowledge of danger signs in a newborn
  - Knowledge of HIV and preventive behaviour
  - Use of modern contraceptives

For the purposes of this survey, a 'household' was defined as a group of persons who shared the same kitchen or hearth. A mother to be considered a member of the household should have lived in the household for at least 6 months. Any household could have representatives from more than one client group. In order to reduce inter-question correlations by sampling more than one sampling universe in the same household, the inclusion criterion stipulated that only one client group could be sampled in a given household. Therefore, interviewers were trained to delineate the composition of the household, and in case of more than one eligible person residing in the household, to select one for the interview using simple random sampling.

2.3. Questionnaire development and interviewer training

Each question in the questionnaire was associated with an indicator in the Network's M&E plan. Questions and option responses were taken from a field-tested Knowledge Practice and Coverage-2000 Instrument [32], which was adapted to local Network conditions. The instrument was translated into Armenian, back translated to English and then pre-tested in a suburb of Yerevan.

Training activities included a two-day training-of-trainers workshop in English, which took place in Yerevan, followed by three simultaneous five-day LQAS pre-survey training workshops in Armenian, which took place in Goris, Gavar and Gyumri. The workshops were supervised by a master trainer who was accompanied by a simultaneous translator from the Network and by members of the Network who were trained as LQAS trainers.

The data collection at each site began immediately following the five-day training and was supervised by a master trainer and trained members of the Network. Data collectors anticipated difficulty in finding eligible respondents due to two major concerns: first, many homes had been abandoned because of emigration; second, the fertility rate in Armenia was low, thereby limiting the number of households having women with babies less than one year old. Regardless of these concerns, the teams finished the data collection on schedule without major problems. Data collection in each SA took approximately 5 days, with two data collectors assigned to each SA.

2.4. Data processing and analysis

With LQAS, data can be analysed in two ways: through data hand tabulation and computer analysis. With hand tabulation, the master trainers lead the data collectors in a process of organizing the raw data taken directly from the questionnaires into tables to ascertain for each SA the number of individuals having the trait of interest. They use the decision rule to identify

indicators, which have reached or not reached the stipulated coverage target. Computer statistical packages can also be used to achieve the same goal after data has been entered into prepared data screens.

While the choice of whether to tabulate data manually or to conduct computer analysis depends on the purpose of the survey, availability of resources, time constraints and the level of details required in the analysis, the two methods complement each other and can be used simultaneously. Each method has its advantages; the main advantage of hand tabulation is that it makes strategic information rapidly available to make urgent programmatic decisions within a couple of days after the survey. If the purpose of the survey is to quickly identify priorities among SAs and to immediately start addressing them, then hand tabulation is an appropriate choice. Computer data entry and analysis take longer to complete, as much as several weeks in many cases. But if the purpose of the survey is to get detailed results, including socio-demographic data, and to conduct complex analyses—for example, a test for association, or even to calculate 95% confidence intervals for each indicator and to weigh the data for differences in SA population size or socio-demographic groups—then a computer database is necessary.

For this survey, both hand tabulations and computer analyses were used. Data hand tabulation was facilitated by the LQAS trainer in each site and conducted by supervisors and data collectors during post-survey workshops. At the end of the workshops, the tables with tabulation results were produced and priority supervision areas for each indicator were identified in each catchment area. Later, data cleaning, computer data entry and analysis were conducted, producing 95% confidence intervals for each indicator and weighting the results for SA population size. Comparison of weighted and unweighted results revealed that the difference between them is small and, therefore, not essential for program planning.

To establish the reliability of the hand-tabulated data, we compared the hand-tabulated results with unweighted computer data. The total error for hand-tabulated data was 3% on average, which is acceptable.

2.5. Identifying program priorities

In the following sections, prevalence measures for key indicators will be presented that measure coverage of the population with health services in the three Network sites.

Network reproductive health program managers used these measures to identify *priority SAs* in their catchment areas, meaning the SAs that fell below average. In the case of ongoing annual monitoring of coverage, a target is used to assess an SA's performance, but as this is the first application of LQAS in Armenia, the average is used as p-Upper so as to establish program targets. The average is also applied to identify SAs that are outliers whose performance is particularly low and below average. When LQAS is done annually, it is the authors' view that using both the average coverage and the coverage target provides more complete information, an approach consistent with identifying SAs that are among the worst of the worst. Once the average coverage is calculated, a data collector uses the Composite LQAS Table (**Appendix 1**) to locate the column header corresponding to the average coverage. If the average is not

http://dx.doi.org/10.5772/66092

divisible by 5, then it is rounded-up to the next highest value, which is divisible by 5 in order to use the LQAS table. Therefore, if, for example, the average coverage was 66%, it would round up to 70%. In the next step, the data collector locates the row for a sample of 19 (or the appropriate sample size if different from 19). At the intersection of this column and row, one finds the decision rule. All decision rules were determined using cumulative probabilities of the binomial model as explained in Section 1. If the total number of correct responses in an SA is less than the decision rule, then the SA is below average or did not reach the target and is in need of special attention. For example, if average coverage for an indicator was found to be 70%, then the decision rule would be 11. Any SA having less than 11 correct responses for that indicator would be judged to be below average.

Appendix 2 demonstrates how data collectors used their data to make judgements about the SAs. It displays summary results for one of the program catchment area, Gyumri. The first indicator in this example is, 'Percentage of mothers attending an antenatal visit by a clinically trained provider.' In this catchment area, there were four SAs. The first series of columns shows the number of correct responses in each of the 4 SAs (13, 10, 13, 16), making the total correct 52. The next series of columns shows that the sample size in each SA was 19, for a total sample size of 76. The average coverage (68%) is calculated and recorded in the far right cell. The Composite LQAS Table was then used by data collectors to determine the decision rule using a rounded-up value of 70, resulting in a decision rule of 11. The highlighted cells indicate those SAs that were found to be below average. Two other indicators are included in this example to demonstrate how these LQAS data were used to identify priorities for local decision-making. The remainder of this chapter uses the aggregate measures only.

3. Results and discussion

This section presents the results of the computer analysis of data from three types of respondents: non-pregnant women aged 15–49 years, men aged 15–54 years, and mothers of children aged 0–11 months. For the sake of simplicity, we refer to these three groups as women, men, and mothers.

The findings are shown for each project site as well as for the Armenia Network for Health as a whole, aggregating the results from the three pilot sites. The computer results are weighted by the population size of each supervision area. Qualitative data presented along with LQAS results complement the findings.

3.1. Family planning

3.1.1. Abortions

As **Table 1** shows, 44% of women reported a history of at least one induced abortion, with a TAR being 2.0 abortions per woman. A higher TAR of 2.6 in ADHS 2000 [33] may be attributed to the fact that our project sites included one small city (Gyumri) and two towns (Goris and Gavar) and did not include the national capital as well as other towns. The abortion rate varied

across the three sites; Goris had a TAR of 1.4 abortions per woman, Gavar's TAR was 2.5, and Gyumri's was about 1.9. The ADHS 2000 results showed that on average an Armenian woman had 50% more abortions than births (2.6 TAR for 1.7 TFR). Further, comparison of the trends in TAR and TFR showed that although the TAR declined over the last 10 years (2.6 in 2000, 1.8 in 2005 and 0.8 in 2010), the total fertility rate remained stable over the same period (1.7), which is below replacement levels. The progressively fewer young woman becoming pregnant and giving birth, together with the overall reduced pregnancy rate [16] over time, may explain, at least in part, the decreasing TAR. Others suggest that with the introduction of medical abortions women practice self-administered abortions and do not necessarily seek medical assistance; as a result the total number of induced abortions is underreported [34]. At this point in its history, Armenia's health strategy needs to encourage and support families to wish to have more children, while also ensuring access to contraceptive methods and information, which will allow other couples to make informed choices and plan their family without an invasive procedure like abortion.

	Pilot areas			
Indicators	**Goris**	**Gyumri**	**Gavar**	**Aggregate coverage**
Abortions				
Percentage of women reporting an induced abortion in lifetime	48% (±11%)	40% (±13%)	43% (±10%)	44% (±6%)
Total abortion rate	1.40	1.85	2.52	2.03
Contraceptive method information and use				
Percentage of mothers counselled about the family planning after delivery	12% (±7%)	25% (±10%)	24% (±9%)	21% (±4%)
Percentage of women who report currently using a family planning method (CPR[a])	61% (±10%)	43% (±11%)	29% (±9%)	42% (±6%)
Percentage of women who report currently using a modern family planning method	16% (±7%)	43% (±11%)	24% (±8%)	26% (±5%)
Post-abortion family planning				
Percentage of women counselled about family planning methods after abortion	29%	45%	39%	37% (±9%)[b]

[a] Contraceptive prevalence rate (CPR) is the percentage of women who are currently using, or whose sexual partner is currently using, at least one method of contraception, regardless of the method used (WHO, http://www.who.int/reproductivehealth/topics/family_planning/contraceptive_prevalence/en/).
[b] CIs for separate sites are not indicated because of small sample sizes.

Table 1. Prevalence and 95%-confidence intervals of induced abortions, total abortion rate, family planning and post-abortion information provision in the three sites of Armenia.

http://dx.doi.org/10.5772/66092

3.1.2. Contraceptive method information and use

The contraceptive prevalence rate (CPRe[3]) was 42%, which is similar to the ADHS 2000 national estimate (39%). However, we detected substantial variation within the Network. Goris had the highest CPR (61%), and Gavar the lowest (29%), while the CPR in Gyumri approximated the Network average at 43% (**Figure 1**). About a quarter of all women used a modern method of contraception; however, there was significant variation among the sites. For example, Goris, the site with the highest CPR, had the smallest proportion of women using modern contraceptive methods (16%), indicating that most women preferred natural methods (e.g., withdrawal, abstinence, rhythm). In contrast, in Gavar, where the CPR was the lowest (29%), 75% of women using any family planning method used a modern method. Gyumri, as the second largest city in Armenia, had the highest rate of modern contraceptive use (43%), while having the second highest CPR among the three sites.

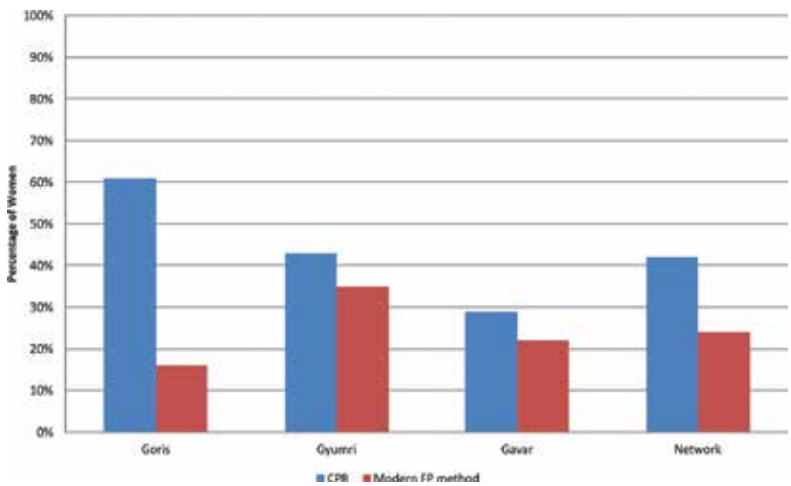

Figure 1. Contraceptive Prevalence Rate and Prevalence of Modern Family Planning Methods in the Three Sites of Armenia.

ADHS results show a slight decrease in the CPR (39% in 2000, 33% in 2005 and 34% in 2010) and variable use of modern contraceptives over a 10-year period (14% in 2000, 12% in 2005 and 17% in 2010) [35].

3.1.3. Post-abortion family planning

Further analysis of the survey results revealed that that only 37% of women who had had at least one induced abortion in their lifetime had been counselled regarding use of contraceptives after their abortion. A similar percentage (32%) was advised to use modern methods after the

[3] Contraceptive prevalence rate is the proportion of women of reproductive age who are using (or whose partner is using) a contraceptive method at a given point in time (WHO, 2006 http://www.who.int/whosis/whostat2006ContraceptivePrevalenceRate.pdf).

abortion, suggesting that those women who received any post-abortion counselling were likely to be informed about modern methods of family planning. One-third of women who had been exposed to post-abortion counselling reported using a modern method of contraception during the survey.

Postpartum experiences were also examined to determine whether mothers received family planning information. Only 21% of mothers reported being counselled about methods of contraception. While service providers play an influential role in the decision to use modern family planning methods in Armenia, some qualitative studies suggest that provider and client satisfaction with family planning counselling are critical factors [36, 37]. It remains unclear if service providers offer sufficient counsel to allay women's fears regarding modern contraceptives and to ensure that they feel empowered to make informed choices to use modern contraceptives.

According to these studies, the primary reasons for abortions were difficult socioeconomic conditions and a desire to postpone or stop childbearing. Among other factors affecting their decisions to abort, women cited insufficient effectiveness of contraceptives, lack of complete information about modern family planning methods, the side effects of the methods, difficulty of access, such as when women had to travel to the district centre to get contraceptives, and a lack of male motivation to use condoms. The views of service providers were also explored. The majority stressed the importance of providing adequate and comprehensive information and the need to discuss with women the variety of methods available. Some health care providers also emphasized the difference in health needs, level of education, preferences of women and respective importance of individual approach.

3.2. Safe motherhood

This section presents findings related to perinatal and newborn care. The respondents were women, men, or mothers, depending on the question. The questionnaire for mothers included both knowledge and behavioural questions, while the ones for women and men were knowledge questions. The outcomes thereof are summarized in **Tables 2** and **3**.

3.2.1. Antenatal care

Mothers with children 0–11 months were questioned about their pregnancy for their infant. Eighty percent of all mothers in the pilot sites said they had visited a clinically trained provider (gynaecologist, doctor, nurse, or midwife) at least once for antenatal care during that pregnancy. The ADHS 2000 found a slightly higher percentage (92%) for this indicator [38]. This variation is likely due to the Network study being confined to three locations, whereas the ADHS is nationally representative.

There was a variation among the sites. Gyumri had a lower percentage (67%) than Gavar (77%) and much lower than Goris (95%). Only 34% of all mothers interviewed had their first antenatal visit during the first trimester. The variation across the three sites for this indicator was similarly substantial (Goris: 45%; Gyumri: 26%; Gavar: 32%).

http://dx.doi.org/10.5772/66092

Indicator	Pilot areas			Aggregate coverage
	Goris	Gyumri	Gavar	
Antenatal care				
Percentage of mothers who visited a clinically trained provider at least once for antenatal care	95% (±5%)	67% (±11%)	77% (±9%)	80% (±5%)
Percentage of mothers who had the first antenatal visit during their first trimester of pregnancy	45% (±10%)	26% (±10%)	32% (±10%)	34% (±6%)
Percentage of mothers receiving iron supplements during pregnancy	6% (±5%)	17% (±9%)	12% (±7%)	12% (±4%)
Knowledge of danger signs				
Percentage of women knowing 2 or more danger signs during pregnancy	64% (±10%)	44% (±11%)	72% (±9%)	63% (±6%)
Percentage of men knowing 2 or more danger signs during pregnancy	33% (±10%)	30% (±11%)	55% (±10%)	42% (±6%)
Percentage of women knowing 2 or more danger signs during labour/delivery	47% (±10%)	43% (±11%)	74% (±9%)	58% (±6%)
Percentage of men knowing 2 or more danger signs during labour/delivery	29% (±9%)	33% (±11%)	44% (±10%)	37% (±6%)
Percentage of women knowing 2 or more postpartum danger signs	69% (±9%)	46% (±11%)	79% (±8%)	68% (±6%)
Percentage of men knowing 2 or more postpartum danger signs	37% (±10%)	41% (±11%)	49% (±%)	44% (±6%)
Percentage of women knowing 2 or more danger signs in each stage	37% (±10%)	31% (±11%)	58% (±10%)	45% (±6%)
Percentage of men knowing 2 or more danger signs in each stage	14% (±7%)	16% (±8%)	33% (±10%)	23% (±5%)

Table 2. Safe Motherhood: maternal care indicators and knowledge of maternal complications with 95%-confidence intervals in the three sites of Armenia.

Overall, a very low proportion of mothers (12%) reported receiving iron supplementation during their recent pregnancy, with Goris (6%) being significantly behind Gyumri and Gavar (17 and 12%, respectively). These results may cause concerns about maternal nutritional status. However, the ADHS 2000 found that only 12% of women in Armenia actually suffer from mild, moderate, or severe anemia [38].

Indicator	Pilot areas			Aggregate coverage
	Goris	Gyumri	Gavar	
Percentage of women knowing 2 or more danger signs in newborns within first 7 days of birth	72% (±9%)	52% (±11%)	78% (±8%)	70% (±6%)
Percentage of men knowing two or more danger signs in newborn with first 7 days of birth	55% (±10%)	42% (±11%)	52% (±10)	50% (±6)

Table 3. Indicators of newborn danger signs awareness with 95%-confidence intervals in the three sites of Armenia.

3.2.2. Knowledge of danger signs

Questions about danger signs during pregnancy, labour/delivery and postpartum were asked of both women and men. Men appeared to be substantially less knowledgeable than women about danger signs at any perinatal stage. Only 63% of women and 42% of men knew two or more danger signs during pregnancy. Gyumri and Goris reported the least knowledge for both men and women (women: 44 and 64%, respectively; men: 30 and 33%, respectively). Fifty-eight percent of women knew at least two danger signs during delivery. Again, Gyumri (43%) and Goris (47%) had lower levels of knowledge than Gavar (74%). Only 37% of men knew at least two danger signs during delivery, with the variation among pilot sites quite similar to the variation among women (**Table 2**).

Knowledge of two or more postpartum danger signs was demonstrated by 68% of women; consistent with other knowledge indicators, Gavar had the highest percentage of knowledgeable women (79%) followed by Goris (69%) and Gyumri (46%). Only 44% of men exhibited knowledge of postpartum danger signs, with similar results across the sites.

Overall knowledge of perinatal complications (knowing two or more danger signs during each of three stages: pregnancy, delivery and postpartum period) was low among both women (45%) and men (23%). Gavar exhibited relatively higher level of knowledge than the other two sites: 58% of the women knew two more danger signs during each of the three stages. Nevertheless, the WHO data show a slow but stable decrease in the maternal mortality ratio in Armenia, from 40 in 2000 and 2005 to 33 in 2010 and 25 in 2015 [39]. This may be due to emergency care in Armenia being accessible or to the increasing awareness of men and women to perinatal danger signs. However, the ADHS 2010 did not include this information and we cannot substantiate our view regarding the declining MMR.

3.2.3. Newborn care

Women and men were asked about newborn danger signs within the first 7 days of birth. Seventy percent of women knew two or more danger signs of a newborn. As with maternal care, the level of knowledge in Gyumri was lower (52%), than in Goris (72%) and Gavar (78%) (**Table 3**).

http://dx.doi.org/10.5772/66092

Among men, 50% knew two or more danger signs in a newborn in the first 7 days of birth. The pattern of variation is similar to that for women, with Gyumri having a lower level of knowledge (42%) than Gavar and Goris (52 and 55%, respectively).

3.3. Breastfeeding and complementary feeding

The questionnaire for mothers of children 0–11 months included questions about initiating breastfeeding, exclusive breastfeeding and introduction of complementary foods (**Table 4**). For some indicators, the data were analysed for subgroups rather than the whole sample. For example, the assessment of *exclusive breastfeeding* included mothers of children ages 0–5 months, while the assessment of *complementary feeding* practices included mothers of children ages 6–9 months. Other indicators included the full sample of women with children 0–11 months.

Indicator	Pilot areas			Aggregate coverage
	Goris	Gyumri	Gavar	
Percentage of mothers of children 0–11 months whose newborns were breastfed within the first hour of delivery	3% (±4%)	31% (±11%)	41% (±10%)	28% (±5%)
Percentage of mothers of children 0–11 months whose newborns were placed with the mother immediately after cutting the umbilical cord	18% (±8%)	42% (±11%)	31% (±10%)	30% (±6%)
Percentage of children 0–5 months exclusively breastfed in the 24 h preceding the survey	22%	24%	22%	23% (±7%)[a]

[a] CIs for separate sites are not indicated because of small sample sizes.

Table 4. Breastfeeding indicators with 95%-confidence intervals among mothers of children 0-11 months in the three sites of Armenia.

Twenty-eight percent of mothers began breastfeeding newborns within 1 h of the birth; values ranged from 3% in Goris to 41% in Gavar. Gyumri, at 31%, was about average. The low percentage of newborns breastfeeding within the first hour of life could be attributed to the practice of separating newborns from their mothers immediately after delivery. Less than one-third of mothers (30%) reported that their babies were given to them right after the cutting of the umbilical cord. Overall, less than a quarter of mothers with children aged 0–5 months (23%) reported exclusive breastfeeding at the time of the survey. In a smaller sub-sample of mothers with children aged 0–3 months, the proportion of mothers who exclusively breastfed was 34%. This result indicated that mothers commenced giving liquids or complementary feeding earlier than recommended.

A trend analysis of exclusive breastfeeding indicates a rapid decline in the practice by the second month of life, and it continues to decline steeply until 4 months of age (**Figure 2**). While two-thirds of mothers (66%) reported they were exclusively breastfeeding their 0–1-month-old babies, the prevalence of exclusive breastfeeding decreased with a child's age, with only 43, 22 and 8% of mothers breastfeeding their babies exclusively in the second, third and fourth months, respectively. By the fifth month, no mothers reported they were exclusively breastfeeding their babies.

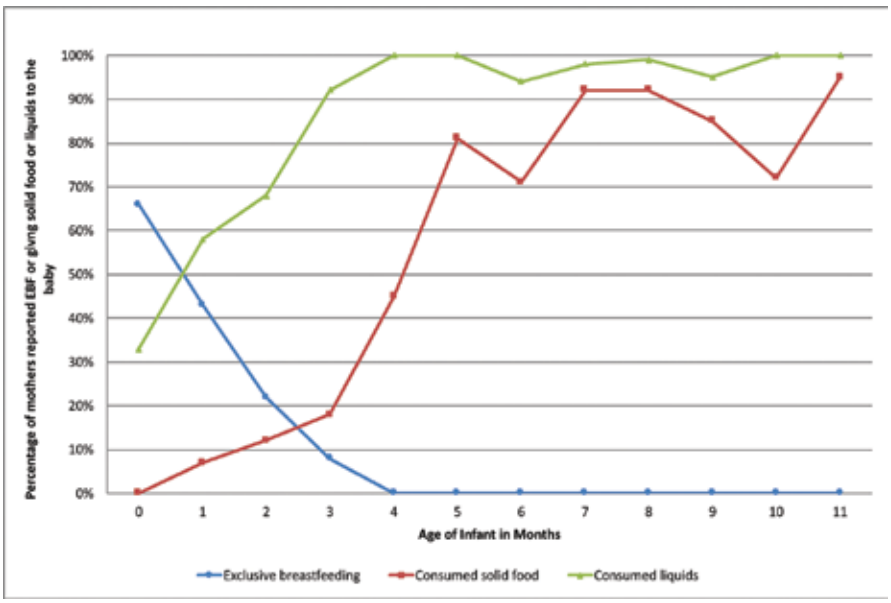

Figure 2. Infant feeding in a cohort of infants aged 0–11 months.

The majority (85%) of mothers of children ages 6–9 months reported that they were giving complementary foods to their babies. More than 90% of mothers in Goris and Gavar fed their 6–9-month-old babies complementary foods, but in Gyumri, only 74% of mothers reported giving complementary foods. Consistently, a high proportion of mothers were giving complementary food to their 6–9-month-old children for each month of age. A trend analysis for complementary feeding showed that 74% of the mothers were practicing complementary feeding at 6 months and 84% by 9 months. When the 'food' category was disaggregated into liquids and solids and all mothers with children ages 0–11 months were included, the earliest premature introduction of solids occurred at 1 month and was 18% by the third month, while 34% of mothers began giving their infants liquids during the first month of life. In ADHS 2010, 23% of mothers of children 2–3 months gave their babies non-milk liquids in addition to breastfeeding, and 14% gave other milk in addition to breastfeeding; 35% of mothers gave complementary foods to their babies of 4–5 months. The tendency to introduce both liquids and solids at an early age, by as early as 1 month, and the consequent rapid decline in exclusive breastfeeding is an area for program action at the present time.

http://dx.doi.org/10.5772/66092

4. Current research priorities in areas of reproductive health in Armenia

In accordance with the statement of the International Conference on Population and Development Program of Action, "reproductive health… implies that people are able to have a responsible, satisfying and safe sex life and that they have the capability to reproduce and the freedom to decide if, when and how often to do so…The aim of interventions is to enhance reproductive health and promote reproductive rights rather than population policies and fertility control" [40].

The results of ADHS 2010 showed that 21% of married women of reproductive age had an unmet need for contraceptives, while the World Bank's data indicated a lower prevalence of 13.5% for the same year [41]. Another source reported this indicator as 19% [42]. In either case, this result indicates that access to modern birth spacing methods and complete information on contraceptives might be still an issue, although there are also other priorities for reproductive health research and policy in Armenia. Infertility and termination of pregnancies due to socio-economic difficulties is now evident and needs further investigation. Fifteen percent of women in ADHS 2010 reported termination of pregnancy due to socio-economic difficulties. Eight percent of women of reproductive age declared being infecund in 2010 as compared to four percent of women in 2005 [43].

Types and causes of infertility need investigation, and prevention and treatment programs should be designed, implemented and evaluated. A law to support families with more than one child should be introduced, its implementation should be tracked, and its effectiveness should be assessed. The low status of women and various socioeconomic barriers contribute to the high level of sex-selective abortions, indicating that programs to empower women and to support families wanting more children need to be implemented and monitored. In all community assessments, the LQAS methodology can be used. Using the modification of LQAS developed for large countries makes it possible to generate statistically rigorous results on both national and regional levels and at an even lower relative low cost as compared to conventional cluster sampling [44].

This observation is not intended to denigrate DHS and MICS, but we believe it is appropriate and preferable to use rapid and affordable methods for recurrent local-level monitoring that supports sound program management. LQAS can be applied not only to survey the community but also to assess health facilities (HFA). Using the LQAS for HFA allows detecting the gaps in quality of health care provision and identifying priorities for resource allocation and systematic monitoring [45–47].

During 2000 when this survey was undertaken, abortion was already an outmoded means of family planning. Today, 15 years later, this is even more true and yet the current TAR in Armenia signals the need for ongoing local level assessments of its prevalence and cause. The results of this study and the methodology used may support the development of sound health policies for Armenia's and other nations.

Appendices

LQAS Table: Decision rules for sample sizes of 12–30 and coverage targets/average of 10–95%

Sample size*	Average coverage (baselines)/annual coverage target (monitoring & evaluation)																	
	10%	**15%**	**20%**	**25%**	**30%**	**35%**	**40%**	**45%**	**50%**	**55%**	**60%**	**65%**	**70%**	**75%**	**80%**	**85%**	**90%**	**95%**
12	N/A	N/A	1	1	2	2	3	4	5	5	6	7	7	8	8	9	10	11
13	N/A	N/A	1	1	2	3	3	4	5	6	6	7	8	8	9	10	11	11
14	N/A	N/A	1	1	2	3	4	4	5	6	7	8	8	9	10	11	11	12
15	N/A	N/A	1	2	2	3	4	5	6	6	7	8	9	10	10	11	12	13
16	N/A	N/A	1	2	2	3	4	5	6	7	8	9	9	10	11	12	13	14
17	N/A	N/A	1	2	2	3	4	5	6	7	8	9	10	11	12	13	14	15
18	N/A	N/A	1	2	2	3	5	6	7	8	9	10	11	11	12	13	14	16
19	N/A	N/A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
20	N/A	N/A	1	2	3	4	5	6	7	8	9	11	12	13	14	15	16	17
21	N/A	N/A	1	2	3	4	5	6	8	9	10	11	12	13	14	16	17	18
22	N/A	N/A	1	2	3	4	5	7	8	9	10	12	13	14	15	16	18	19
23	N/A	N/A	1	2	3	4	6	7	8	10	11	12	13	14	16	17	18	20
24	N/A	N/A	1	2	3	4	6	7	9	10	11	13	14	15	16	18	19	21
25	N/A	1	2	2	4	5	6	8	9	10	12	13	14	16	17	18	20	21
26	N/A	1	2	3	4	5	6	8	9	11	12	14	15	16	18	19	21	22
27	N/A	1	2	3	4	5	7	8	10	11	13	14	15	17	18	20	21	23
28	N/A	1	2	3	4	5	7	8	10	12	13	15	16	18	19	21	22	24
29	N/A	1	2	3	4	5	7	9	10	12	13	15	17	18	20	21	23	25
30	N/A	1	2	3	4	5	7	9	11	12	14	16	17	19	20	22	24	26

N/A: *not applicable*, meaning LQAS cannot be used in this assessment because the coverage is either too.

unshaded cells indicate where alpha or beta are < 10%
shaded cells indicate where *alpha* or *beta* errors are ≥ 10%.
hashed cells indicate where *alpha* or *beta* errors are > 15%.

Appendix 1: Composite LQAS Table.

Summary Results: Baseline Survey, December 2000												
Indicator	**Number of correct responses in each SA**				**Total number of correct responses**	**Sample size in each SA**				**Total sample size**	**Average coverage**
	Decision rule										
	1	**2**	**3**	**4**		**1**	**2**	**3**	**4**		
Section 3A: Prenatal care											
Percentage of mothers at least once visited a clinically trained provider for antenatal care	13	10	13	16	52	19	19	19	19	76	68%
	11	11	11	11							
Section 3B: Delivery and newborn care											
Percentage of mothers of children 0–11 months whose newborns were placed with mother immediately after cutting umbilical cord	15	5	5	12	37	19	19	19	19	76	49%
	7	7	7	7							
Section 4: Family planning											
Percentage of mothers whose most recent birth was planned	17	18	12	16	63	19	19	19	19	76	83%
	14	14	14	14							

Appendix 2. A sample of hand tabulated data—summary results for mothers with children 0–11 months—baseline survey.

Author details

Joseph J. Valadez* and Lusine Mirzoyan

*Address all correspondence to: joseph.valadez@lstmed.ac.uk

Liverpool School of Tropical Medicine, Liverpool, UK

References

[1] Family planning/Contraception: Fact sheet N°351. World Health Organization, WHO, 2015. http://who.int/mediacentre/factsheets/fs351/en/

[2] Facts on induced abortions worldwide: Fact sheet. Guttmacher Institute, USA. May, 2016. http://www.who.int/reproductivehealth/publications/unsafe_abortion/abortion_facts/en/

[3] Say L, Chou D, Gemmill A, Tunçalp Ö, Moller A, Daniels J, Gülmezoglu M, Temmerman M, Alkema L.Global causes of maternal death: a WHO systematic analysis. The Lancet. 2014;2(6):e323–e333.

[4] Black R, Laxminarayan R, Temmerman M, Walker N. Reproductive, Maternal, Newborn, and Child Health. Disease Control Priorities, Washington, DC, World Bank, 2016. Third Edition (Volume 2): Chapter 3.

[5] Ethical issues in obstetrics and gynecology. FIGO Committee for the Study of Ethical Aspects of Human Reproduction and Women's Health; 2012.

[6] Gissler M, Kauppila R, Meriläinen J, Toukomaa H, Hemminki E. Pregnancy-associated deaths in Finland 1987–1994: definition problems and benefits of record linkage. Acta Obsetricia et Gynecolgica Scandinavica. 1997;76:651-657.

[7] Reardon D, Coleman P. Short and long term mortality rates associated with first pregnancy outcome: population register based study for Denmark 1980–2004. Medical Science Monitor. 2012;18(9):PH71-6.

[8] Parazzini F, Vecchia C, Negri E, Checchetti G, Fedele I. Reproductive factors and the risk of invasive and intraepithelial cervical neoplasia. British Journal of Cancer. 1989;59:805-809.

[9] Remennick L. Induced abortion as cancer risk factor: a review of epidemiological evidence. Journal of Epidemiology and Community Health. 1990;44:259-264.

[10] Weiss N. Events of reproductive life and the incidence of epithelial ovarian cancer. American Journal of Epidemiology. 1983;117(2):128-139.

[11] Vecchia C, Negri E, Franceschi S, D'Avanzo B. Reproductive factors and the risk of hepatocellular carcinoma in women. International Journal of Cancer. 1992;52:351.

[12] Induced abortion operations and their early sequelae. Journal of the Royal College of General Practitioners.1985;35(73):175-180.

[13] Freedman M. Comparison of complication rates in first trimester abortions performed by physician assistants and physicians. American Journal of Public Health. 1986:76(5): 550-554.

[14] OECD. Development in Eastern Europe and the South Caucasus: Armenia, Azerbaijan, Georgia, Republic of Moldova and Ukraine. OECD Publishing, 2011. http://dx.doi.org/10.1787/9789264113039-en

[15] Armenia Demographic and Health Survey. 2000. National Statistical Service Armenia, Ministry of Health Armenia, ORC Macro, USA.

[16] Armenia Demographic and Health Survey. 2010. National Statistical Service Armenia, Ministry of Health Armenia, ICF International, USA. 2011.

[17] Report: Prevalence of and reasons for sex-selected abortions in Armenia. UNFPA, Armenia. 2012.

[18] The baby doom: selective abortions in Armenia, The Armenian Weekly; November 23, 2011.

[19] Moughamyan T. In Armenia, Abortion Rates are High and Access to Contraception is Limited. 2012. Available from: http://www.ourbodiesourselves.org/2013/02/in-armenia-abortion-rates-are-high-and-access-to-contraception-is-limited/[accessed: 2016-07-16]

[20] Valadez J. Commentary: learning to be creative with HIV/AIDS studies: looking for the variation—not only the average. International Journal of Epidemiology. 2009;38(1): 214-216.

[21] Valadez J, Devkota B, Pradhan M, Meherda P, Sonal G, Dhariwal A, Davis R. Improving malaria treatment and prevention in India by aiding district managers to manage their programs with local information: a trial assessing the impact of Lot Quality Assurance Sampling on program outcomes. Tropical Medicine & International Health. 2014;19(10):1226-1236.

[22] Turner A, Magnani R, Shuaib M. A not quite as quick but much cleaner alternative to the Expanded Program on Immunization (EPI) Cluster Survey design. International Journal of Epidemiology. 1996;25(1):198-203.

[23] Pagano M, Valadez JJ. Understanding practical lot quality assurance sampling. International Journal of Epidemiology. 2010;39(1):69-71.

[24] Anoke S, et al. Comparing two survey methods of measuring health-related indicators: Lot Quality Assurance Sampling and Demographic Health Surveys. Tropical Medicine & International Health, 2015;20(12):1756-1770.

[25] Valadez J. Assessing Child Survival Programs in Developing Countries: Testing Lot Quality Assurance Sampling. Cambridge: Harvard University Press; 1991.

[26] Valadez J, Weiss W, Leburg C, Davis R. Assessing community health programs: A Trainer's guide. Using LQAS for baseline surveys and regular monitoring. 2nd edition. Teaching-aids At Low Cost, St Albans; 2007.

[27] Robertson S, Valadez J. Global review of health care surveys using lot quality assurance sampling (LQAS), 1984–2004. Social Science & Medicine. 2006;63:1648-1660.

[28] Hund L, Northrop-Clewes C, Nazario R, Suleymanova D, Mirzoyan L, Irisova M, Valadez J. A novel approach to evaluating the iron and folate status of women of reproductive age in Uzbekistan after 3 years of flour fortification with micronutrients. PLoS One. 2013;8(11):e79726.

[29] Robertson SE, Anker M, Roisin AJ, Macklai N, Engstrom K. The lot quality technique: a global review of applications in the assessment of health services and diseases surveillance. World Health Statistical Quarterly. 1997;50:199-209.

[30] Reinke WA. Industrial Sampling Plans: Prospects for Public Health Applications. Occasional Paper. Baltimore: Institute for International Programs, The Johns Hopkins University School of Hygiene and Public Health, 1988.

[31] Turner AG, Magnani RJ, Shuaib M. A not quite as quick but much cleaner alternative to the Expanded Program on Immunization (EPI) Cluster Survey design. International Journal of Epidemiology. 1996;25(1):198-203.

[32] CORE Group. Knowledge, Practice and Coverage Survey –2000+. Washington, DC, October, CORE Group, 2000. Available from: http://pdf.usaid.gov/pdf_docs/Pnack209.pdf.

[33] Saribekyan K, Abrahamyan R, Balasanyan M, Hovhannisyan A. Maternal and child health. National Statistical Service of the RA. 2000; Chapter 12: p 125.

[34] Jilozian A, Agadjanian V. Is induced abortion really declining in Armenia? Studies in Family Planning. 2016;47(2):163-78. doi:10.1111/j.1728-4465.2016.00053.x.

[35] Westoff C. A new approach to estimating abortion rates: DHS analytical studies 13. Macro International, USA; 2008, p 8.

[36] Thompson M, Harutyunyan T, Ghukasyan G: Feasibility Study: The Strategic Introduction of the Standard Days Method of Family Planning in Armenia: Formative Research Final Report. Yerevan, Armenia: American University of Armenia, The Center for Health Services Research and Georgetown University, Institute for Reproductive Health; 2001.

[37] Salvador S, Danielian L. Report on Qualitative Research: JHU/PCS Project on Reproductive Health in Armenia. Yerevan, Armenia: American University in Armenia, The Center for Health Services Research and The Center for Policy Analysis; 1999.

[38] Reproductive, maternal and child health in Eastern Europe and Eurasia: a comparative report. CDC, ORC Macro DHS. Atlanta, USA. 2003; Chapter 12: p 156,157.

[39] Maternal Mortality in 1990–2015: Armenia. WHO, UNICEF, UNFPA, World Bank Group, and United Nations Population Division Maternal Mortality Estimation Inter-Agency Group. Geneva, World Health Organization: 2015. Available from: http://www.who.int/gho/maternal_health/countries/arm.pdf [accessed 2016-08-15]

[40] Guidelines on Reproductive Health. United Nations Population Information Network (POPIN). Available from: http://www.un.org/popin/unfpa/taskforce/guide/iatfreph.gdl.html [accessed 2016-07-05]

[41] The World Bank Group. 2016. Available from: http://data.worldbank.org/indicator/SP.UWT.TFRT?locations=AM [accessed 2016-09-03].

[42] Alcema L, Kantorova V, Menozzi C, Biddlecom A. National, regional, and global rates and trends in contraceptive prevalence and unmet need for family planning between 1990 and 2015: a systematic and comprehensive analysis. The Lancet. 2013;381(9878): 1647. Available from: http://dx.doi.org/10.1016/S0140-6736(12)62204-1 [accessed 2016-09-03].

[43] Armenia Demographic and Health Survey. 2005. National Statistical Service Armenia, Ministry of Health Armenia, ORC Macro, USA. 2006.

[44] Hedt B, Olives C, Pagano M, Valadez J. Large Country-Lot Quality Assurance Sampling: A New Method for Rapid Monitoring and Evaluation of Health, Nutrition and Population Programs at Subnational Levels. Washington, DC: World Bank Group; 2008.

[45] Berendes S, Lako RL, Whitson D, Gould S, Valadez JJ. Assessing the quality of care in a new nation: South Sudan's first national health facility assessment. Tropical Medicine & International Health. 2014;19(10):1237-1248.

[46] Oladele EA, Ormond L, Adeyemi O, Patrick D, Okoh F, Oresanya OB, et al. Tracking the quality of care for sick children using lot quality assurance sampling: targeting improvements of health services in Jigawa, Nigeria. PLoS One. 2012;7(9):e44319.

[47] Valadez JJ, Transgrud R, Mbugua M, Smith T. Assessing family planning service-delivery skills in Kenya. Studies in Family Planning. 1997;28(2):143-150.

Section 2

Quality Frameworks and Management

Chapter 5

A Framework to Manage Quality of Enterprise Content Management Systems

José González Enríquez,
Francisco José Domínguez Mayo,
Julián Alberto García García,
María José Escalona Cuaresma and
Manuel Mejías Risoto

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66199

Abstract

There is a wide range of enterprise content management (ECM) systems which supports, among other things, document management processes, records management and Web content management. However, each of these systems has many features and some of them can meet organizational needs depending on the scale, sector and workflow of the organization. In addition, it is very common that organizations are unaware of what ECM system best fits their needs, since each company has its particular scope and strategic objectives. This chapter is contextualized within the real project called THOT designed for the Andalusian Public Administration in Spain. The aim of this project is to study in detail ECM systems and propose an objective method to compare them for the specific scope and strategic objective of organizations. Quality evaluation framework (QuEF) has been adapted for this purpose.

Keywords: software engineering, standards, methodologies, enterprise content management, quality analysis and evaluation

1. Introduction

Today's world economic situation is dominated by concepts such as globalization, which involves the relocation of companies, the constant search for lower costs to maximize profits

by continuous mergers and acquisitions, and a constant motivation for improving and optimizing all processes. Defining business models to manage organizations effectively is essential and nowadays, it has become a common practice followed by a large number of companies in all areas of business, especially within the information and communications technologies (ICT) context [1]. This can be extrapolated to the context of document processes, as these can help sales and product development organizations capitalize on new business opportunities [2]. For example, this implementation would permit a faster commercialization of new offerings, a flexible response to customers' needs, a quicker reply to changing market dynamics and business competitiveness improvement.

Every day, organizations have to handle a lot of different type of information such as general documents, audio-visual files/resources, management reports, customer cards or invoices, among others. In addition, this information is not disconnected, but integrated within the organization's business process. The ever-growing use of this kind of information within ICT organizations confirms this tendency for the years to come. There are two main alternatives to manage efficiently and effectively these organizations' information resources. The first one consists in specifying, designing and implementing an ad-hoc system for this goal within organizations. This alternative is not usually feasible in most of them due to the high cost of maintenance, development and future updates.

The second option deals with using proprietary or open source enterprise content management (ECM) solutions. Performed literature review indicates that this option best fits companies' budgets.

A variety of ECM definitions can be found in the literature, according to the Association for Information and Image Management (AIIM) [3]. The ECM consists of strategies, methods and systems used to capture, manage, store, preserve and deliver content and documents related to organizational processes. ECM systems (ECMSs) and strategies facilitate an organization's unstructured information management. However, there currently is a wide range of document management solutions with different cost and functional scope (such as interoperability, security, usability, efficiency, customization among others). Therefore, to choose the one which best meet the organization's necessities is not a simple task. Smith and McKeen [4] define ECM solutions as an integrated approach to manage all organization's information including paper documents, data, reports, Web pages and digital assets. In addition, Tyrväinen et al. [5] provide a framework to stimulate and guide future research as well as point out research issues specific to the ECM field. The aforementioned authors argue that ECM offers an important and complex subfield of information systems. As regards the term ECM, these authors add that it has been widely adopted by software product vendors and practitioners in order to refer to technologies used to manage the content of assets like documents, Web sites, intranets and extranets in organizational or inter-organizational contexts. Preliminary findings suggest that ECM has not managed to attract adequate attention.

Nordheim et al. [6] also examine the topic of ECM solutions, but under a strategic development and implementation process for a large oil company. The authors represent a case of a hybrid development approach to ECM that involves the lifecycle, teleological and dialectical engines of development. They state that this is opposite to the evolutionary development motor, which

http://dx.doi.org/10.5772/66199

has prevailed in the hitherto reported content management research. Nordheim et al. also show a case study which complements process-based research on enterprise system implementations in general. In addition, they suggest that research and practice on large-scale ECM implementations should acknowledge all the four motors of change.

Furthermore, Munkvoldet et al. [7] explain that the concept of ECM represents integrated enterprise-wide management of the lifecycles of all forms of recorded information content and their metadata, organized according to corporate taxonomies and supported by appropriate technological and administrative infrastructures. This publication is based on a case study about a Norwegian oil company (Statoil), where they identify a wide range of issues related to content, infrastructure and change management. As ECM perspective is concerned, Munkvoldet et al. argue that ECM perspective is found to integrate and extend the existing research areas of information resource and document management, as well as the repository model of knowledge management. Thus, they show how ECM deserves further attention beyond its current market hype, as a potential area of Information Systems (IS) research crossing several previously separate areas of information management from the enterprise viewpoint.

All ECM definitions show that there is a wide range of ECM descriptions and solutions in literature which support, among other things, document management processes, record management and Web content management. The problem is that each of these solutions has many features and some of them can meet organizational needs depending on the scale, sector and workflow of the organization. In addition, it is very common that organizations are unaware of what ECM system best fits their needs because each company has its particular scope and strategic objectives.

This chapter is contextualized within the THOT project, which is a real project designed for the Andalusian Public Administration in Spain. Document processes management is essential and critical in this context, since e-Government is taking a key role in setting the strategic plans of the Public Administration. Therefore, the AOPJA (Public Agency of Contracting Services for Transport and Infrastructure Constructions) and the University of Seville are executing the THOT project as a result of these needs. THOT is an innovative project with high costs (621,250 Euros) focused on document management applied to service agreement records and transport infrastructure projects. Consequently, deciding on the most suitable ECM solution for the AOPJA context poses considerable responsibility. An ECM system that meets all requirements must be developed once the most suitable ECM solution is characterized and defined for the AOPJA context.

Therefore, THOT project aims to analyse in detail strategies and document management systems to investigate and define an innovative solution, so that records management can improve. Then, the objective is to study ECM systems in detail and propose an objective method to compare them for the specific scope and strategic objective of the AOPJA.

For this purpose, a quality evaluation framework (named QuEF) is proposed to analyse and evaluate ECM solutions. QuEF is a work-in-progress framework that has been used in a case of study (Public Administration of the Regional Government of Andalusia, Spain) in order to

validate it. Moreover, QuEF provides an agile, flexible and efficient solution based on a Web environment, allowing organizations to choose the best ECM system for their necessities and ensure the continuous quality improvement of these systems in the organization.

For these reasons, the objective of this chapter is twofold: to identify assessment criteria for ECM solutions, focused on the AOPJA context, by means of the proposed framework and to demonstrate the value QuEF offers to support technology acquisition.

As such, this work is structured as follows: Section 2 summarizes some of the most recent work related to ECM systems and reference standards on this topic.

Section 3 describes QuEF and its theoretical foundations and Section 4 presents the THOT project.

Then, Section 5 explains how the framework has been applied to the project and introduces each phase of the framework and its execution process along the project.
Finally, Section 6 states learned lessons and on-going work.

2. Related work

There are different versions of ECM systems found in literature; Scott [8] evaluates the factors that lead to users choosing an ECM system. The results show the importance of cognitive involvement with technology highlighting the importance of including cognitive participation in building acceptance studies.

Alalwan and Weistroffer [9] performed a comprehensive literature review of ECM. A conceptual framework of the areas of interest in relation to ECM as well as an agenda for future research on this topic were proposed.

Ninety-one ECM publications were reviewed in this work. The authors concluded that the ECM systems involve interacting technical, social, organizational and business aspects. Also, the authors suggested that the current ECM literature could be grouped into three main research pylons:

- The first one consists of the four ECM component dimensions.
- The second one deals with the enterprise system lifecycle.
- The final one constitutes the strategic managerial aspect. An agenda for future based on the review and the suggested conceptual framework is also suggested.

Smith and McKeen include the strategies, tools, processes and skills an organization needs to manage its information assets on ECM solutions along its lifecycle. The authors explained that an effective ECM strategy should address each of the four lifecycles stages:

1. Capture: gathering all activities associated with collecting content.
2. Organize: indexing, classifying and linking content and databases together to provide access within and across business units and functions.

http://dx.doi.org/10.5772/66199

3. Process: shifting and analysing content in such a way that may facilitate.

4. Maintain: ensuring that content is regularly updated.

The authors pointed out that most ECM initiatives take a bottom-up approach that focuses on delivering immediate benefits through projects such as intranet portals, information searching and Web content management, while the top-down vision for ECM includes improved decision-making, better utilization of information and collection of competitive intelligence. However, performed research indicated that managers also acknowledge the fact that greater value can be gained from taking a more strategic approach to ECM. The authors showed that those organizations that can effectively "hand-shake" content stewardship practices with appropriate information behaviours, and values and information technology on a broader scale, can have a significant effect on their performance.

Grahlmann et al. [10] explain that ECM centers on managing all types of content used in organizations. The authors present an overview of previous research explaining that scientific literature on ECM is limited and no consensus on the definition of ECM is reached. Therefore, the literature review surfaces several ECM definitions that are merged herein into a more consistent and comprehensive definition of ECM. Indeed, the authors mentioned above provide the functional ECM framework (FEF) which is an overview of the potential functionalities of ECMSs. They applied FEF to three case studies to communicate on ECMSs, to familiarize oneself with and to direct future research. It may also form the basis for more formal reference architecture, and practitioners may use it as an assessment tool for comparing the functionalities provided by existing ECMSs.

Herbst et al. [11] showed that ECM is an important enabler of informatioey identified a set of critical success factors for ECM and develop a framework that helps organizations assess their readiness for ECM. In Herbst et al., this framework was developed following the data collected in workshops held between ECM project leaders and members of five companies. The authors argued that expert's opinions and experiences are combined with research results from the academic literature, and two illustrative cases showed how the framework has been put into practice.

In addition, Rickenberg et al. [12] explained that ECM can be considered an integrated approach to information management. They exposed that ECM research is still an emerging field of IS research, even though practitioners pay much attention to this concept. Furthermore, they provided a detailed review of the body of academic research: the ECM domain, its evolution and the characterization of the main topics. In Rickenberg et al., an established ECM research framework is adopted, refined and explained together with its associated elements and working definitions. On this basis, 68 publications were reviewed and classified, and concepts were derived. Prior research was synthesized and findings were integrated in a concept-centric way. Finally, the authors exposed implications for research and practice, including future trends.

Two works are related to the ECM implementations. Haug [13] included a definition of a process model for ECM implementation in SMEs. This author proposed a new pattern definition for ECM technology development. In [14], Van Rooij explained that legal issues

generated by ERP could be similar to those generated by ECM systems. Therefore, it is advised, with appropriate adaptation, to take these issues into account when developing strategies for implementing the ECMs.

ISO 2709:2008 [15] specifies the requirements for a generalized exchange format containing records describing all forms of material capable of bibliographic description as well as other types of records. ISO 15836: 2009 [16] establishes a standard for describing resources across domains known as metadata elements Dublin Core Set. This standard defines the elements that are commonly used in the context of an application profile, limiting their use in accordance with the policies of a particular community and it does not define the implementation details.

ISO 10244:2010 [17] provides businesses with the tools to identify the relevant aspects of the business work processes and document them in a standardized format.

3. The THOT project scenario

The THOT project is an e-Government project with the objective to implement an ECM system in the Public Administration of Andalusian region of Spain granted for 621,250.00Euros. This chapter is focused on the first phase of the THOT project and explains how the technological and functional status of existing ECM systems has been studied. It is very important to evaluate all existing alternatives in the market in order to align the scope of the organization with its purpose. It is also relevant not to reject the decision because it is difficult to change the chosen system due to the cost, once the development of the solution has started. As the evaluation process concerns, a static evaluation or characterization of an ECM solution is not enough, since new improvements of ECM systems are continuously appearing and one has to compare alternatives dynamically and objectively. In addition, this work considers the different preferences of the elements containing an ECM system in terms of given specific context. This study discusses in detail existing ECM systems in the market and proposes an objective method to compare them within the specific scope and strategic objective of organizations.

Nowadays, the Andalusian Public Administration is driving the need for a change in the following document management systems:

- JUPITER [18], although this situation will change, because this information system it is going to be replaced by an ERP technology platform
- TheERIS-G3 [19], which is an e-Government system to manage the electronic procedures to process public records in each public agency
- @rchivA [20], which facilitates the Patrimony Documentation management of Andalusia

This project aims to cover different disciplines of research and innovation as document management, electronic government, dissemination and Web services integration policies, enabling organizations to provide a common framework for document management. To achieve the aforementioned objectives, the project is being carried out along the following stages and activities:

http://dx.doi.org/10.5772/66199

1. *Studying technological and functional status*
   - Analysing the current situation
   - Benchmarking existing tools and systems and selecting the most appropriate option
   - Defining new document management functionalities and fitness for certification as a document management system
2. *Setting the context of research results*
   - Adjusting Andalusian horizontal documentary series
   - Developing policies to preserve digital documents
3. *Defining the solution*
   - Defining the document management solution
   - Developing and implementing a basic functional system
   - Defining a dissemination system
4. *Disseminating results*
   - Defining the project dissemination plan
   - Defining project dissemination indicators
   - Executing the dissemination plan

In this context, QuEF is proposed as a work-in-progress framework to be validated and used for the evaluation of existing ECM systems. This framework offers a suitable methodology to analyse and evaluate ECM solutions dynamically and objectively. In addition, it includes methods to calculate preferences of ECM features and it defines a lifecycle and tool support to enforce the quality continual improvement of an ECM solution.

4. The QuEF framework methodology

QuEF [21] is a work-in-progress framework that has been used in this case study (Public Administration of the Regional Government of Andalusia, Spain) in order to validate it. QuEF is a framework to manage quality of entities (products, processes, services or organizations, among others) in any context and domain. In previous work, this framework was used to manage quality in model-driven Web development methodologies. In addition, this framework can also be used for consumers to identify the most suitable product or process for them and decide accordingly.

This framework describes templates to define a specific Quality Model for the domain under study. It also offers a method to customize the Quality Model, evaluate it and calculate the preferences of its elements. Besides, the framework includes the definition of a set of phases

to enforce continuous quality improvements in the Quality Model. The most important aspect is that quality management is the central quality entity in the Quality Model. Furthermore, a tool support is also implemented in order to promote this solution in real environments. Therefore, users may indeed obtain quality management in an automatic way using QuEF, and hence, computerise the quality management of entities (products, processes, services or organizations, among others) with the final aim to reduce cost and time and improve overall quality. As such, this framework provides

- A set of phases to enforce the continuous quality improvement
- Quality standards and international best practices
- Methods for each phase and templates to customize the Quality Model
- Multi-criteria methods to calculate the elements preference value of the Quality Model
- A tool to support the quality management lifecycle process

Table 1 shows the relationship among other standards. It also represents the relation of these standards as well as the best practices and approaches that have been applied to the QuEF framework both to define QuEF itself and to apply it to a specific domain.

Standards, best practices and approaches	Work application
ISO/IEC 20000, ITIL	ISO/IEC 2000 standard and ITIL best practices deal with improving service quality based on a quality continual improvement of the service lifecycle. For instance, ITIL defines Strategy phase, Design phase, Operation phase, Transition phase and quality continual improvement phase. QuEF covers the same idea with a different goal since QuEF framework manages quality-based on a quality continual improvement of the Quality Model lifecycle.
TQM, Six Sigma, CMMI Planguage, C-INCAMI or CTQ, among others	The QuEF framework defines different phases with artifacts, methods and tools for each phase. Most of these approaches could be adapted and applied to some phases of QuEF. They cover the similar aspects between QuEF and quality management Strategy and Operation phase. For instance: • TQM is a management integrative philosophy aims at continuously improving the quality of products and processes. It could be applied to Strategy and QCI phases in QuEF. • Six Sigma is a business process management strategy very similar to TQM working with many established quality-management tools. Most of them could be used in Strategy phase and Operation phase of QuEF. • CMMI is a process improvement approach that intends to help organizations improve their performance. Therefore, it could be applied to Strategy and quality continual improvement phases. • Planguage could be applied to the Strategy phase of QuEF for specifying quality. • C-INCAMI provides a domain (ontological) model defining all the concepts and relationships needed to design and implement processes. Hence, it could be used in a Strategy and Operation phases. • CTQ could also be applied to the Strategy phase, the Design phase and the Operation phase of QuEF for specifying project context, nonfunctional requirements, measurement, evaluation and analysis.

Table 1. Standards, best practices and approaches related to the QuEF framework.

http://dx.doi.org/10.5772/66199

As **Figure 1** shows, the framework can be used from two points of view: providers', who need to analyse, control, evaluate and improve entities and consumers, who need to compare entities (depending on their context) to decide on the most suitable one for them.

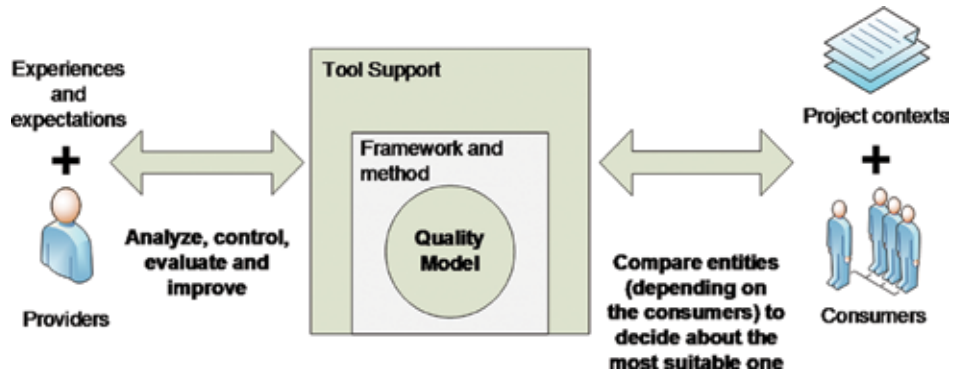

Figure 1. Conceptual scheme representing the goals to be achieved with QuEF.

It mainly differs from other frameworks in that it focuses on the Quality Model as well as defines a lifecycle where all phases turn around that Quality Model, as shown in **Figure 2**.

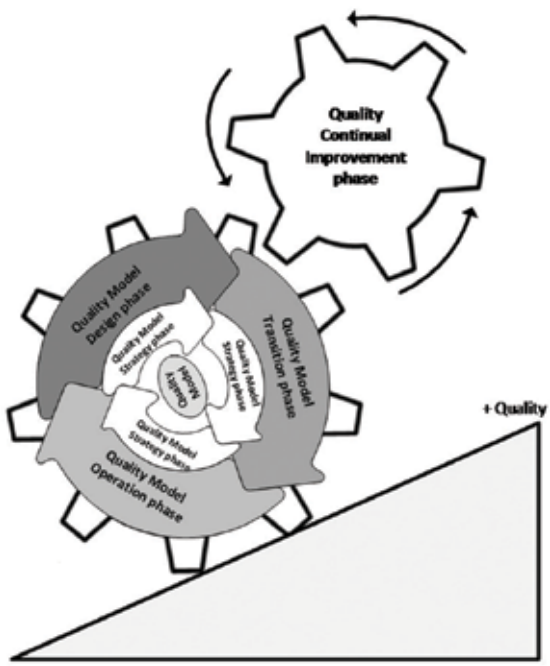

Figure 2. Quality management based on the Quality Model lifecycle.

Moreover, QuEF provides an agile, flexible and efficient solution based on a Web environment, so that organizations can choose the most suitable ECM system for their purposes as well as enforce the continuous quality improvement of these systems within the organization. It is based on ITIL v3, but with a difference; it does not focus on services, but on a Quality Model. Similarly to ITIL v3, it comprises of five phases to ensure the continuous quality improvement of the Quality Model. The aim is to centralize all quality management efforts on the Quality Model. This means that it incorporates several phases including different objectives and artefacts. The aforementioned phases are

- Quality Model Strategy (QMS) phase: This phase is strategic active that focuses on the definition of a quality management strategy. The past, present and future view elements of the Quality Model in the domain under study are essential to achieve an effective and efficient quality management process.
- Quality Model Design (QMD) phase: This is the phase where the Quality Model is finally designed depending on the requirements from the previous phase. It is the model used in the next phase for the quality management performance.
- Quality Model Operation (QMO) phase: In this phase, the Quality Model is used to carry out the quality management process. Consequently, the analysis and evaluation management processes are performed within this phase.
- Quality Model Transition (QMT) phase: This phase describes the processes that execute changes in the Quality Model, in cases where the domain or context changes due to the appearance of new trends, but without affecting the Operation phase.
- Continuous Quality Improvement (CQI) phase: This phase performs all mechanisms to improve quality in all processes in the lifecycle and the Quality Model.

An effective and efficient quality management essentially demands to define the domain under study. Thus, it is important to consider what type of ECM system is concerned. It is not the same to develop an ECM system for a bank, where security stands as a more important quality characteristic, as opposed to the development of an ECM system as presented herein, where usability, functionality and performance are crucial. The purpose of QuEF is not only to assure a clear strategy for quality management but also automatically facilitates a continuous quality improvement by means of generating checklists and documentation, as well as automatic evaluations and plans which control and improve quality and thus, automatically, reduce effort and time.

Figure 3 shows the specific proposed metamodel for QuEF. There are many definitions in literature trying to clarify what a Quality Model is. In QuEF, it means a set of characteristics and its relationships, which constitutes the base to specify quality requirements and evaluate them. The Quality Model represents its core with quality management revolving around it. This work proposes a Quality Model metamodel consisting of a simplification and adaptation of the ISO/IEC 15939, so that the model implementation can be more flexible and practical. The Quality Model contains Features, Sub-Feature and Property.

http://dx.doi.org/10.5772/66199

- Feature (FT-<Level 1>): It is a general concept that involves a set of higher-level concept of ECM system properties that describes it. It includes a set of Sub-Features.
- Sub-Feature (SF-<Level 0>): It includes a specific concept of an entity. It is a set of lower-level concept properties of an entity. It is also utilized to categorize ECM systems in two levels (Feature and Sub-Feature).
- Property: It indicates the degree to which a Sub-Feature is measured by the use of a Metric. Particularly, a property is used for describing and analysing the Sub-Features of an entity. It is an element of an ECM system. In other words, a property is used for describing and analysing Sub-Features.

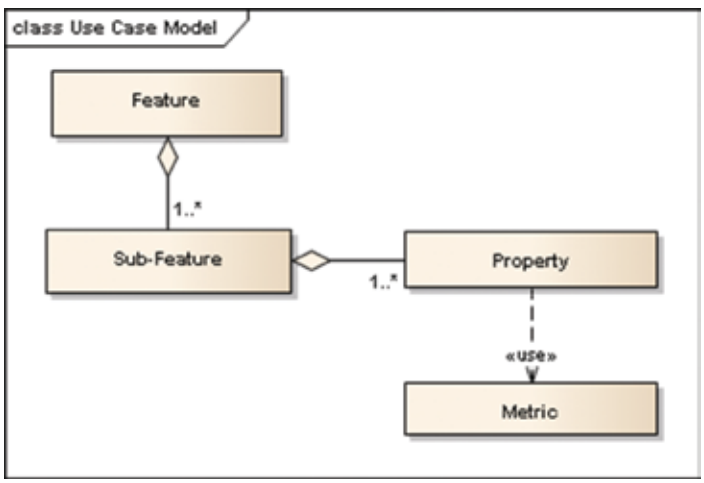

Figure 3. Quality metamodel on QuEF.

4.1. Case study: applying the QuEF framework to the THOT project scenario

As previously mentioned, the THOT project is a transfer project carried out in collaboration with the Regional Government of Andalusia (Junta de Andalucía) in Spain. It has two main objectives:

1. to obtain a detailed analysis and evaluation of ECM systems applied to contracting records for infrastructure projects and, to find out and
2. to define an innovative solution that may improve procedural records management.

Currently, there are several solutions for this type of systems in the market, although herein, the most appropriate ones have been selected in relation to the scope of the project. Thus, the QuEF framework has been used to obtain the detailed analysis and evaluation of different ECM alternatives. As **Figure 4** shows, two points of view have been identified in the THOT project scenario: ECM system providers who need to analyse, control, evaluate and improve ECM systems and the Public Administration of the Regional Government of Andalusia, which needs

to compare ECM systems (depending on the THOT project scenario) to make a decision on the most suitable one to apply.

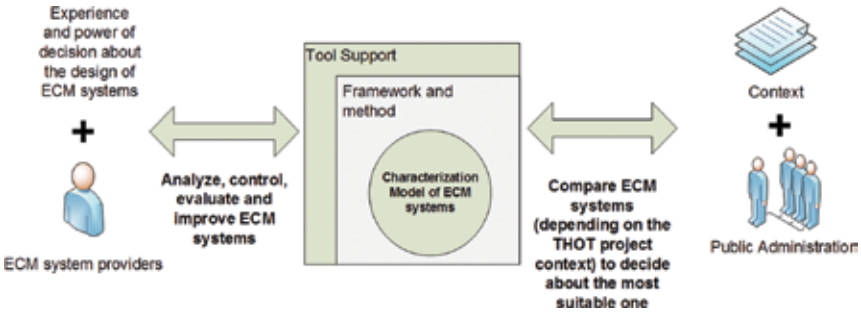

Figure 4. Conceptual scheme representing the application of QuEF framework to the THOT project scenario.

In order to apply the QuEF framework, IWT2 research group (Web Engineering & Early Testing) is developing QuEF-TS with the aim of automating all processes and artefacts that QuEF defines for each phase. In this particular case, an Enterprise Architect (EA) modelling environment with UML 2.2 is provided.

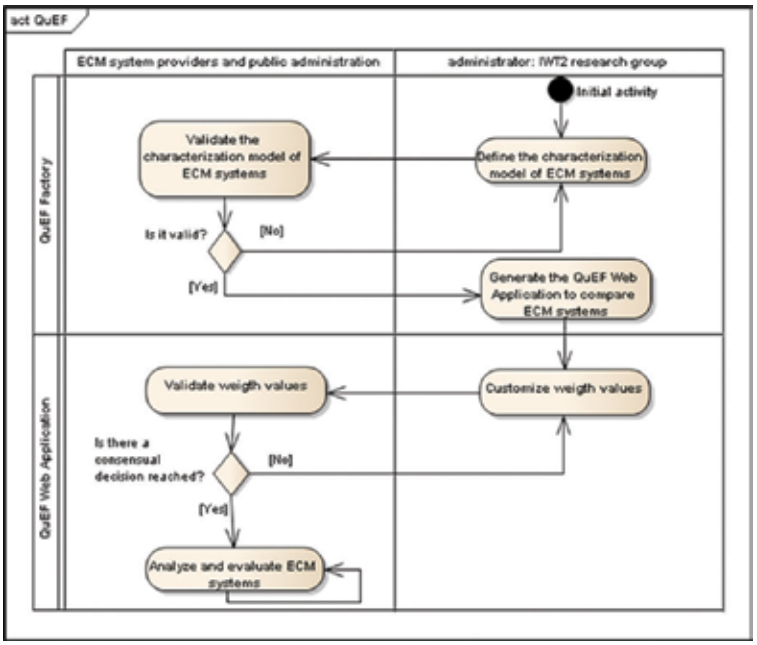

Figure 5. Activity diagram for the tool support use.

Activity diagrams are used to describe the business and operational step-by-step workflows of components within the system. **Figure 5** shows the overall flow of control between ECM

http://dx.doi.org/10.5772/66199

system providers and Public Administration may be observed (not necessarily registered users) as well as the system administrator (in this case, the IWT2 research group).

Then, the Quality Model lifecycle has been applied to obtain a characterization model of ECM systems. Firstly, the Strategy phase has been established for an effective quality management of ECM systems. In this case, the quality management problem is, in fact, a multiple-criteria decision-making or multiple-criteria decision-analysis (MCDA) problem, because multiple criteria are used to calculate the preferences regarding the Quality Model elements demand. Besides, it is a multi-objective optimization problem, as it is important not only to implement the most valuable ECM system but also to reduce cost, risk and uncertainty.

Secondly, the Design phase has been applied to fix the Quality Model that was used to analyse and evaluate quality. The transition phase is also important, since new systems or trends regarding the insertion of new characteristics in the Quality Model can appear in future iterations and these changes must be controlled.

Thirdly, this study will explain how the ECM systems have been analysed and evaluated in the Operation phase. Hence, a tool support has been implemented to automate the generation of all artefacts.

Finally, the CQI phase is shown to clarify how this in conjunction with its phases have been set in order to perform an improved quality cycle encapsulating all phases within the framework.

4.2. The Quality Model Strategy phase

In this phase, it is important for the THOT project purposes to explain that the characterization model must be based on concrete solutions for ECM systems and not in work associated with theoretical proposals in the ECM systems context. The main motivation of this decision is the fact that a new ECM system must be developed, which will be based on an existing ECM system. In addition, the high cost this type of projects entails makes it very risky to implement a solution from the ground up. Thus, it is better that all expected requirements and functionalities have already been validated in the market.

Then, a characterization model has to be defined in order to analyse and evaluate the systems with QuEF. Carrying out this task demands that the main Features and Sub-Features of all these systems are known. A systematic literature review (SLR) was used to analyse the current situation. A SLR is a means of identifying, evaluating and interpreting all available documents related to a particular thesis in a specific investigation area.

To perform the SLR, the protocol defined by Kitchenmham [22, 23] was chosen. It is one of the most acknowledged in software engineering. This model establishes the necessity of specifying some research questions (RQs) that will guide the work. For this work, the following RQs were proposed:

- RQ1 - What ECM systems currently exist in the market and what do they offer?
- RQ2 - How can ECM systems be adapted to the general guidelines of the Andalusian Public Administration?

- RQ3 - What is the most appropriate ECM system that the Andalusian Public Administration, and more specifically, the contracting services for transport and infrastructure constructions must use?
- RQ4 - What areas of improvement are needed for the selected ECM system?

The databases considered for this SLR were ACM Digital Library, EiCompendex, IEEE Xplore, ISI Web of Knowledge, Science Direct, SCOPUS, Springer Link and Wiley InterScience Journal Finder.

Once the method was applied, the results showed that the following tools needed to be analysed: Alfresco [24], Documentum [25], Nuexo/Athento [26], IBM FileNet [27] and OpenText [28]. Then, the concept mapping method [29, 30] was executed in order to obtain the characterization model. This method involves all stakeholders in the project. Several meetings were organized with all stakeholders and system providers to discuss requirements in their systems and all this information was used to build the characterization model by the concept mapping method. Concept mapping is a general method that can be used to help any individual or group describe their ideas about some topics in a pictorial form.

In accordance with all these analysed systems and the strategy followed herein, a set of preferences for each element of the Quality Model was defined for adapting it to the project scope.

4.3. The Quality Model Design phase

The QMD phase in QuEF is understood to encompass all the relevant elements to design the Quality Model. As such, the QMD phase defines the necessary basic characteristics to be analysed using the defined templates. Defining these characteristics, it is possible to asses each solution uniformly.

The characterization scheme is composed of 10 features, which respond to the questions identified in the QMS phase. This priority is contextualized within the needs of this project: to define an innovative solution for document management applied to procurement of services and transport infrastructure projects within the Regional Government of Andalusia, Spain. However, the software allows any potential user utilizing QuEF to set priorities based on his/her own needs. These basic characteristics are illustrated below:

FT01: Functional modules. The results obtained in the QMS phase point out that a valid ECM system must include natively and minimally the following functionalities.

FT02: User orientation. Although ECM systems offer standard solutions on its orientation towards the end user, many companies need to use easy and versatile systems because not all their employees have the same user profile to handle computer tools.

FT03: Functionality to capture, access, retrieve and view documents. The ability that lets anyone transform the system depending on the organization preferences or user profile.

http://dx.doi.org/10.5772/66199

FT04: Documental lifecycle. This feature enables the user to assess the level or degree of support the system offers to the document cycle. The following Sub-Features of this Feature are described as follows:

FT05: Workflows. This Feature assesses whether the tool supports management with business processes.

FT06: eGovernment. These Features measure the degree of support offered in the context of access and use digital documents.

FT07: Interoperability compliance. A specific section dealing with interoperability has been included: Integration with tools. This Sub-Feature evaluates whether the ECM system provides mechanisms (e.g. APIs) to integrate with third-party tools.

FT08: Security and control. One of the major objectives of document management solutions is to ensure information security. This is facilitated by controlling access to the system from inside and outside the organization and managing the relevant documents in such a way that they are either archived or destroyed. Consequently, these solutions must provide services that ensure that the information stored is secure. It evaluates whether the system is functional enough to analyze data, or otherwise, whether the system allows using third-party tools.

FT09: Architecture. It evaluates whether the system has an open or closed architecture.

FT10: Cost. Cost (both initial and long-term by maintenance) is one of the most important factors any organization must take into account when choosing an ECM solution.

FT11: Assistance and RM (Roadmap) support. This last Feature listed in the latter group includes aspects for the evaluation of the characteristics support, assistance and roadmap provided by the ECM solution.

4.4. The Quality Model Transition phase

The QMT phase provides guidance to undergo changes in the Quality Model without affecting the QMO phase. It helps to know how to handle changes in the Quality Model.

Along this first iteration of the framework, lots of new trends were considered likely to be included in the Quality Model. This study considered in the beginning Alfresco, Documentum, IBM FileNet and OpenText. Then, in a second iteration of the framework, Nuxeo/Athento and KM (SAP 2013) were included. Nevertheless, KM was rejected in the QMS phase because this system did not comply with the outlined project's scope, thus Nuxeo/Athento was finally considered. New Features and Sub-Features were included in the Quality Model and fixed in the QMD phase.

4.5. The Quality Model Operation phase

This phase provides guidance to analyse, evaluate and plan the CQI of ECM systems. In the THOT project, the decision concerning the suitability of the most appropriate ECM system was taken that was deemed most suitable for it. As a result thereof, the Quality Model and each set of preferences pertaining thereto have been defined. As such, this weighted Quality Model

need to be used so as to analyse and evaluate the different systems. Thus, in this phase, the model was implemented to manage quality in ECM systems, which were analysed by means of checklists. These checklists are artefacts that contain all Features, Sub-Features and Properties that have been defined to analyse an entity. Hence, checklists are used in order to know the current state of an ECM system.

In addition, QuEF factory has been developed to automate all tasks in this phase. It allows the user to generate all set of artefacts in each phase of QuEF.

Henceforth, the QuEF factory is being developed as a plug-in of EA. It means that one has to define the Quality Model in EA and explain in which directory the generation of the QuEF-O Web application ought to be created (**Figure 6**).

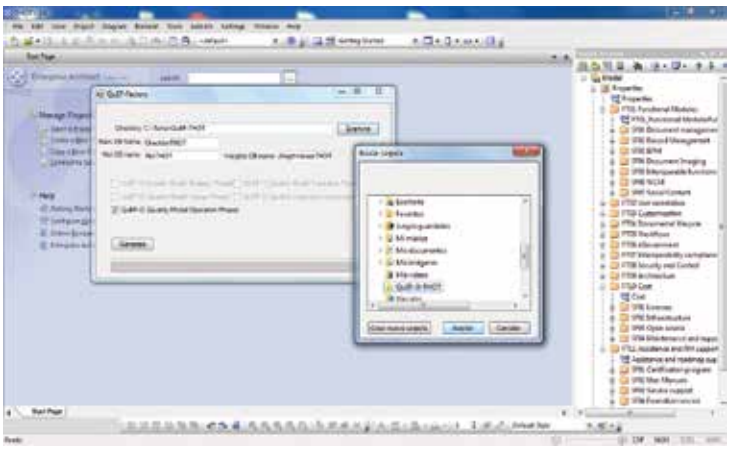

Figure 6. QuEF factory.

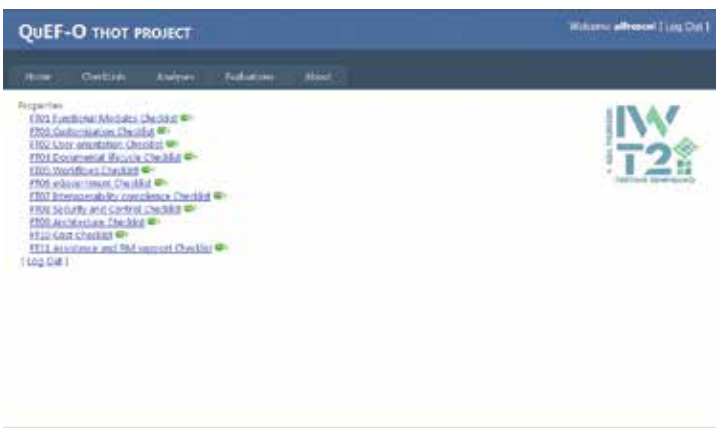

Figure 7. Checklist menu in the QuEF-O Web application lifecycle.

http://dx.doi.org/10.5772/66199

Then, the QuEF-O Web application is generated and the user can use all necessary artefacts to analyse and evaluate all ECM systems found. For instance, **Figure 7** shows all checklists that have been generated in terms of the Quality Model that have been defined.

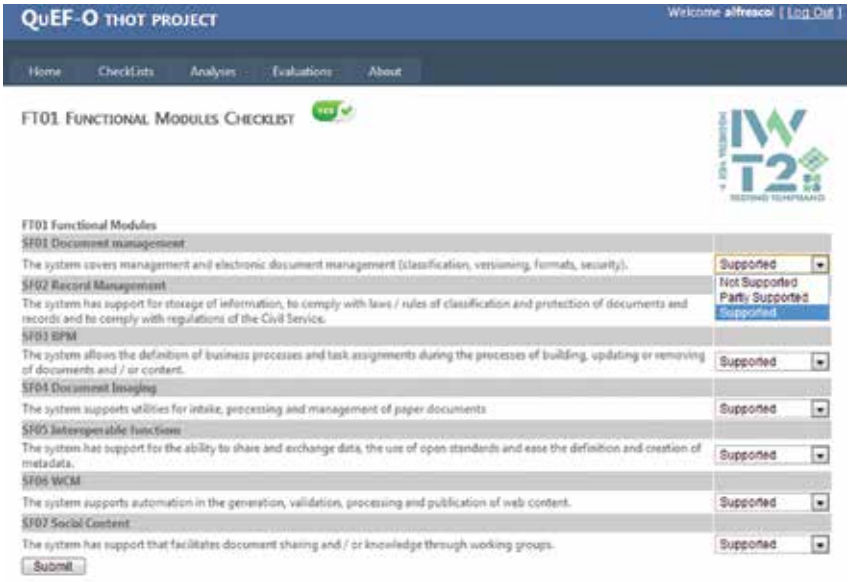

Figure 8. Functional modules checklist in the QuEF-O Web application.

Figure 9. Defined preferences for the sub-features of Functional modules of the QuEF-O Web application.

Figure 8 represents the checklist to analyse the functional modules Feature and Sub-Features. This set of checklist is used in order to study all ECM systems.

As far as preferences are concerned, these have to beset for each element of the Quality Model. For instance, **Figure 9** shows the preferences defined in the QMS phase for functional modules Feature.

Finally, each ECM system can be evaluated in terms of the project's scope (**Figure 10**), taking into consideration the preferences of each element. Consequently, the user can decide what ECM system is the most suitable for his/her needs. Thus, if new elements have to be included in the evaluation, then new iterations of the framework have to be carried out and all artefacts are generated automatically using the QuEF factory. It is the author's view that this way may lead to cost reduction, effort and time associated with and may even improve quality in the quality management processes.

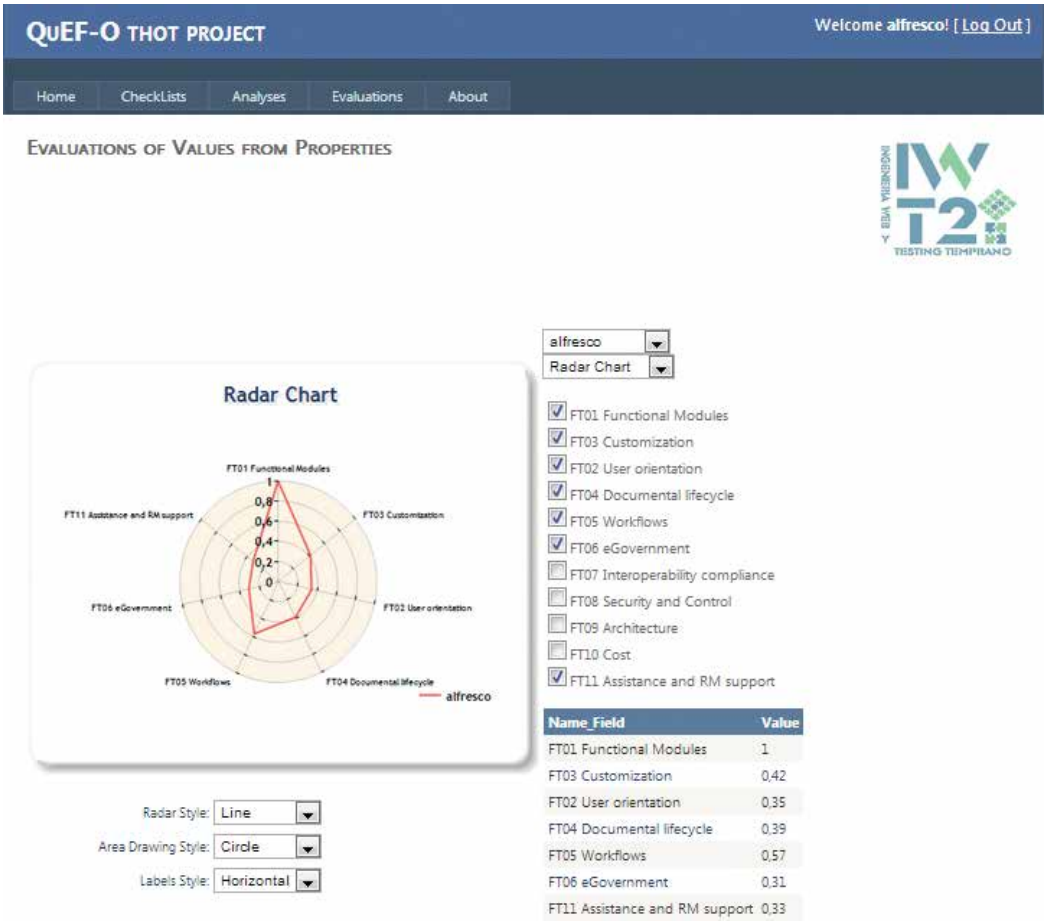

Figure 10. QuEF-O Web application evaluation interface.

http://dx.doi.org/10.5772/66199

4.6. Quality Continual Improvement phase

It is very important to consider the continuous quality improvement in a quality management process based on the Quality Model lifecycle. Thus, this objective can be achieved only through a constant monitoring and measurement of all activities and processes involved in quality management. The main goals of this phase are summarized as follows:

- To use methods from quality management in order to learn from past successes and failures
- To recommend improvements to all processes and activities involved in the Quality Model management
- To control and analyse the basic characteristics as well as monitor and validate them in real-life cases (environments)
- To suggest improvements to increase Return on Investment (ROI) and Value on Investment (VOI) associated with properties
- To support the Strategy and Design phases for the definition of new needs and basic characteristics or processes/activities associated with them
- The results of this phase of the lifecycle must incorporate all the necessary information to
  - Improve the quality of the Quality Model provided
  - Add new properties and basic characteristics that best fit users' properties and the market
  - Improve and streamline internal processes of the approaches

Therefore, it is the authors' view that all these practices along the project ought to be considered so as to achieve a continuous quality improvement encompassing all phases of the framework.

5. Conclusions and future work

This chapter shows the results of a research project that aim to implement an ECM system in the Public Administration and has been carried out in a real environment.

The SLR carried out focused on Features and Sub-Features of ECM solutions that were implemented in existing systems. This is because in the THOT project, the conceptualization of ECM systems is not as relevant as the definition of a set of Features and Sub-Features to compare existing ECM systems in the market. The conceptualization of ECM systems constitutes a detailed research to guide providers of these systems. Therefore, getting a common Quality Model including the requirements that all these systems must fulfil in the future is essential.

Nevertheless, this project offers providers and users the opportunity to decide what existing ECM system in the market is the most appropriate for their purpose. This decision is of paramount importance, as it is going to be used to implement all business requirements in Public Administration. In consequence, the high cost prevents changing the development

project of the new solution, once started. In other words, the project does not focus on the development of a new ECM system, but on deciding which is the most appropriate one in the market.

This study has proposed an evaluation framework concerning different alternatives of ECM systems, and choosing the most suitable one for the scope and context thereof, with the final aim of warranting continuous quality improvement.

QuEF is a framework to manage quality in any product or process, so it can be applied to any entity. It is based on ITIL v3 with the difference that it does not focus on services, but on a Quality Model. Similarly to ITIL v3, it comprises of five phases to ensure the continuous quality improvement of the Quality Model. The aim is to centralize all efforts of quality management on the Quality Model. In addition, the framework also defines protocols and methods to perform each phase, so that all protocols and methods can be systematized. Besides, QuEF-Factory is a tool support that can generate a Web application for each phase in terms of a Quality Model (QuEF-S, QuEF-D, QuEF-T, QuEF-O and QuEF-QCI). QuEF and its tools improve quality management effectiveness and efficiency, since it clarifies the purposes and objectives of management.

Moreover, a set of tools have been evaluated regarding different aspects. Studying the results, they show that Alfresco System and Nuxeo/Athento should be discarded because of the cost and the functionality offered, respectively. The final user should choose between IBM FileNet and Documentum ECM because both are very similar in terms of cost/value.

Acknowledgements

This research has been partially supported by the NDTQ-Framework project (TIC-5789) of Junta de Andalucía, by the TEMPROS project of the Spain Ministry (TIN2010-20057-C03-02), by the SoftPLM network, TIN2015-71938-REDT, the FEDER Funds and Fujitsu Laboratories of Europe (FLE). We also thank all the staff and researches of the Agency of Public Works of the Regional Government of Andalusia, Spain, for their help, support and professional attitude.

Author details

José González Enríquez*, Francisco José Domínguez Mayo, Julián Alberto García García, María José Escalona Cuaresma and Manuel Mejías Risoto

*Address all correspondence to: jose.gonzalez@iwt2.org

Web Engineering and Early Testing (IWT2) Research Group, Computer Languages and System Department, University of Seville, Seville, Spain

http://dx.doi.org/10.5772/66199

References

[1] Mohamed Z. (1997). Business process management: A boundaryless approach to modern competitiveness. Business Process Management Journal, 3(1), 64–80.

[2] IDC White Paper sponsored by Ricoh. (2012). Organizational Blind Spot: The Role of Document-Driven Business Processes in Driving Top-Line Growth.

[3] AIIM. (2013). The Global Community of Information Professionals, http://www.aiim.org/, Accessed 4 May 2016.

[4] Smith, H. A., & McKeen, J. D. (2003). Developments in practice VIII: Enterprise content management. The Communications of the Association for Information Systems, 11(1), 41.

[5] Tyrväinen, P., Päivärinta, T., Salminen, A., & Iivari, J. (2006). Characterizing the evolving research on enterprise content management. European Journal of Information Systems, 15(6), 627–634.

[6] Nordheim, S., & Päivärinta, T. (2006). Implementing enterprise content management: from evolution through strategy to contradictions out-of-the-box. European Journal of Information Systems, 15(6), 648–662.

[7] Munkvold, B. E., Päivärinta, T., Hodne, A. K., & Stangeland, E. (2006). Contemporary issues of enterprise content management. Scandinavian Journal of Information Systems, 18(2), 4.

[8] Scott, J. E. (2011). User Perceptions of an Enterprise Content Management System, 44th Hawaii International Conference on System Sciences (HICSS), p.1,9, 4–7 doi: 10.1109/HICSS.2011.473

[9] Alalwan, J. A., & Weistroffer, H. R. (2012). Enterprise content management research: a comprehensive review. Journal of Enterprise Information Management, 25(5), 441–461.

[10] Grahlmann, K. R., Helms, R. W., Hilhorst, C., Brinkkemper, S., & van Amerongen, S. (2012). Reviewing enterprise content management: A functional framework. European Journal of Information Systems, 21(3), 268–286.

[11] Herbst, A., Simons, A., vomBrocke, J., & Derungs, R. (2014). Critical Success Factors in Enterprise Content Management: Toward a Framework for Readiness Assessment. In Enterprise Content Management in Information Systems Research (pp. 109–124). Springer Berlin Heidelberg.

[12] Rickenberg T. A., Neumann M., Hohler B., & Breitner M., Enterprise Content Management - A Literature Review. AMCIS 2012, available online at http://aisel.aisnet.org/amcis2012/proceedings/DataInfoQuality/10, Accessed 29 July 2012.

[13] Haug A., (2012). The implementation of enterprise content management systems in SMEs. Journal of Enterprise Information Management, 25(4), 349–372.

[14] Van Rooij, J. C. (2013). Legacy Issues in the Implementation of Enterprise Content Management (ECM). International Journal of Information, 3(3), 120–123.

[15] ISO2709:2008, Information and documentation—format for information exchange, http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41319, Accessed 4 July 2016.

[16] ISO15836:2009, Information and documentation—the Dublin Core metadata element set, http://www.iso.org/iso/catalogue_detail.htm?csnumber=52142, Accessed 4 July 2016.

[17] ISO10244:2010, Document management—business process baselining and analysis, http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=45935, Accessed 4 July 2016.

[18] JUPITER Project, Government of Andalucia, http://www.juntadeandalucia.es/repositorio/ usuario/listado/fichacompleta.jsf?idProyecto=19, Accessed 15 July 2016.

[19] ERIS-G3 Project, Government of Andalucia, http://www.juntadeandalucia.es/repositorio/ usuario/listado/fichacompleta.jsf?idProyecto=680, Accessed 15 July 2016.

[20] Archiva (@rchiva), Government of Andalucia, https://ws024.juntadeandalucia.es/ae/admine-lec/areatecnica/archiva, Accessed 4 July 2013.

[21] Domínguez-Mayo F.J., Escalona M. J., Mejías M., Ross M., & Staples G. 2012. A Quality Management Based on the Quality Model Lifecycle. Computer Standards & Interfaces. 34(4), 396–412.

[22] Kitchenham et al. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. Version 2.3. Department of Computer Science, University of Durham, Durham, UK. EBSE-2007-01.

[23] Kitchenham B., Brereton O. P., Budgen D., Turner M., Bailey J., & Linkman S. (2009). Systematic literature reviews in software engineering—A systematic literature review, Information and Software Technology, 51(1), 7–15, ISSN 0950-5849, http://dx.doi.org/10.1016/j.infsof.2008.09.009

[24] Alfresco, available online at http://www.alfresco.com., Accessed 19 July 2016.

[25] Documentum, available online at http://www.emc.com/domains/documentum/index.htm, Accessed 19 July 2016.

[26] Nuxeo/Athento, available online at http://www.athento.com/nuxeo/, Accessed 19 July 2016.

[27] IBM FileNet, available online at http://www-03.ibm.com/software/products/us/en/filecontmana/, Accessed 19 July 2016.

http://dx.doi.org/10.5772/66199

[28] OpenText, available online at http://www.opentext.es/, Accessed 19 July 2016.

[29] Social research methods, http://www.socialresearchmethods.net/kb/conmap.php, Accessed 17 November 2014.

[30] Trochim, W.M.K. (1989). An introduction to concept mapping for planning and evaluation. Evaluation and Program Planning, 12, 1–16.

Chapter 6

Exploring the Relationship of Supply Chain Risk Management to Quality Management

Tyler Florio

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/65847

Abstract

This research explores the relationship between an organization's supply chain risk management (SCRM) maturity and quality maturity. SCRM maturity was measured using a survey questionnaire sent to organizations in the USA. Quality maturity was assessed via ISO 9001:2008 certification status as well as through a survey questionnaire of total quality management (TQM) practices for organizations in the USA. The results suggest that ISO 9001:2008 is not related to SCRM maturity, while TQM maturity is related to SCRM maturity. Organizations with more mature TQM programs appear to also have more mature SCRM programs.

Keywords: quality management, supply chain management, risk management

1. Introduction

Maintaining the integrity of a supply chain through risk mitigation is crucial to smooth and efficient business operations. However, as supply chains become more global in scope, the potential for risk events occurring increases. For this reason, supply chain risk management has gained substantial interest in recent years among academics. Preliminary research indicates that there are no established standards for certifying risk management capability of organizations in the supply chain. In relation to standards and guidelines, the International Organization for Standardization (ISO) has emerged as a body that seeks to establish and promote best business practices through certifications that organizations can earn. These certifications signal to potential partners that a level of capability has been attained by the organization in a specific area of interest, such as quality management.

Common certifications related to business include, but are not limited to: ISO 9001, 11000, 14001, and 22000. Most of these involve quality and safety, but none certify for supply chain risk management specifically. The closest standard to achieving this is the ISO 31000 risk management principles and guidelines. However, a certification is not available. Therefore, the rigor of a certification is absent from this standard. In order to be awarded an ISO certification, an organization must submit an application to ISO and undergo a rigorous six-stage evaluation process based on various criteria. Other certification bodies do exist that attempt to augment risk management activities in the supply chain. One standard that at least implies an organizational ability to manage risk is Customs-Trade Partnership Against Terrorism (C-TPAT). It is a voluntary program that focuses on improving the security of private companies' supply chains with respect to terrorism. However, it addresses only one particular type of risk event related to disruptions produced by actual or potential terrorist threats and does not address an organization's overall risk management maturity. Nonetheless, organizational certification by ISO 9001 in particular may be able to signal risk management capabilities simply by virtue of the attention the standards bring to improving process management through the principles of total quality management (TQM). ISO 9001 was revised in 2000 to incorporate the principles of TQM into its certification criteria. Therefore, by extension, TQM maturity, in general, may also provide a signal of supply chain risk management maturity.

A large number of organizations, covering a wide variety of industries, are ISO 9001 certified. Because quality standards and certifications are intended to unify and improve business practices as a whole, the following question arises: are companies that have more mature quality systems, and certified to ISO 9001 in particular, better equipped to manage risk? Researching companies who have quality management programs and how their processes have improved since implementing them may shed light on protecting a company's assets, operations, and its structure from adverse risk events. This research assists in confirming the following statement: A company's use of a TQM system, and particularly through ISO 9001 certification, ensures a high level of risk maturity as compared to that of companies that do not implement a TQM system and/or quality certifications.

In the next section, the literature is reviewed and research questions stated. This is followed by the research methodology, which includes a description of the data. Results are then presented. The findings of this research are subsequently discussed and conclusions drawn. This is followed by a discussion of areas for further study and limitations to the research.

2. Literature review

A study involving supply chain risk starts with a classification of all potential risks. These typically include the following: supply risk, demand risk, process risk, technology risk, logistics risk, information risk, and environment risk [1, 2]. The current research focuses on supply and demand risks because these present supply chain managers with significant challenges due to the severity of the impact and difficulty of effective mitigation. Van Miegham [3] characterizes the loss of a key supplier as having a high effect for aggregate loss severity

http://dx.doi.org/10.5772/65847

and a moderate probability of occurrence. Further, Chen et al. [4] report that demand risks have a direct negative effect on supply chain performance. The literature also recommends differing approaches to moderating the occurrence of supply or demand disruptions, such as firm innovativeness, process modularity, and interactive complexity reduction [5–7]. These all relate to the types of activities involved in a quality management strategy that seeks to simplify, standardize, and generally improve products and processes.

As noted, the occurrence of a supply chain risk event can be damaging to any organization no matter where they may be within the supply chain. Technically speaking, risk is defined as the (negative impact to objectives × likelihood of occurrence). Risk management contains four primary steps within its processes. These steps include the following:

- Risk identification,
- Qualitative or quantitative assessment,
- Risk prioritization, and
- Response planning and risk monitoring.

Within these four steps are the proper responses to various types of risk. The first possible response is to mitigate risk. Mitigating entails performing an action to reduce the impact or likelihood of various risk events. The second possible response is to avoid risk. Avoiding risk entails completely ceasing the various activities that create said risk. The third response is to transfer risk. Transferring risk involves shifting risk to other operational areas of the supply chain that are better equipped to handle risk events. The fourth and final way to manage risk is to accept risk. Risk may be accepted if its consequences do not outweigh the benefits of surrounding the risk that is created [8]. Three common methods of assessing risk are effective, but not unified in their approach.

The first method is the Delphi method. The Delphi method was originally developed in 500 B.C. by Greek prophets [9]. The prophets would hear various people's complaints, develop a response, and allow the people to formulate a revised complaint. This method was revised by the RAND Corporation in the 1950s and followed the Greek's original method. The RAND Corporation's method consisted of surveys, followed by a response, followed by revised surveys based on initial results. This process gave participants a chance to re-assess criteria and re-evaluate based on the responses, which provided greater insight to their issues at hand. The main drawback to this method is it is time consuming, because of the analysis of the initial surveys followed by the revisions of the second round of surveys.

The second popular method is Monte Carlo Simulation. This method was developed in the Monte Carlo casinos to gauge risk brought upon by the various gambler's chances of winning. This method focuses on uncertain risk and is assessed by model construction and analysis through computer simulation. The negative aspect of this method occurs, because it requires extensive mathematical prowess and requires significant amount of education to be proficient.

The third method is decision tree analysis. Decision tree analysis utilizes graphical methods to draw correlations to common risks. This method is effective, because it creates a visual image of how various risks are linked, but suffers because of its simplicity [10].

Normally, since avoiding risk is so vital, there should be a series of guidelines in place to facilitate conducting supply chain risk management (SCRM) while supporting total quality management (TQM). Performed literature review indicates that there is no unified series of standards of avoiding risk. Even though most companies have their own emergency preparedness plans in place, a majority of company executives do not review or approve them and only 42% conduct emergency preparedness practices on a regular basis [11]. Companies must learn how to handle diverse amounts of risk such as natural disasters, political unrest, or acts of terrorism. Prior to September 11, 2001, preparedness levels addressing terrorism did not exist.

The World Economic Forum (WEF) has reported that although supply chain risk is an important issue, it is widely mismanaged [12]. Consistent mismanagement of risk across multiple industries might have a ripple effect on global risk which tends to amplify the disruptive impacts of a local risk event with resulting impacts far beyond the corporate sector. The believed cause is companies' inability to detect these ripples (such as the effects of catastrophes and pandemics on the other side of the world) before they become waves that disrupt their supply chains [13]. Many of these preparedness plans have been found to be inefficient because of lack of communication and collaboration, such as Alabama and Louisiana's responses to Hurricane Katrina [14].

Along with natural disasters and man-made risk, there is organizational and network risk. Organizational risks include inventory risk, process/operational risk, quality risk, and management risk, while network risks result from interactions between organizations within the supply chain. Agility, flexibility, contingency planning, and preparedness are preferred generic strategies for managing such risks in general [15]. It has been found that larger companies in the private sector are better equipped to individually handle disaster than smaller companies in the public sector [16–18]. These small companies simply do not have the resources to develop a proactive approach, so instead they take a more defensive (reactive) approach constituting risk elimination aspects [19]. Performed research indicates a gap between the risk management policies that large companies are capable of implementing and what their smaller counterparts are able to produce. Many have tried to develop supply chain risk management policies linking risk identification, risk assessment, and risk mitigation to risk performance [20]. Academics along with industry leaders see a need to replace traditional and varying risk management techniques for ones that are better designed to handle extreme complexities, unpredictable events, and threats. They have sought to discover a link between vulnerable factors and controllable factors. The Supply Chain Resilience Assessment and Management (SCRAM) framework was developed, but only served to suit the needs of a select few companies [21]. Companies must also have a framework in place to address general risk between inter-organizational partnering. Once two or more companies become partners, they assume the other's risk in some or a large capacity, as previously mentioned [22]. According to Zhao et al. [23], "*…supplier, internal, and customer integration are the most important drivers for schedule attainment, competitive performance, and customer satisfaction, respectively"* (p. 115). They find that supply chain risks are negatively related to supply chain integration. Although global

http://dx.doi.org/10.5772/65847

integration is crucial and necessary to be competitive, it inherently carries an increased amount of risk.

Performed literature review shows that multiple organizations and companies have tried to institute best practices for mitigating risk. However, there is not one general set standard for how to approach the subject of avoiding or correcting risk in the supply chain. The reason for this variety in practices is possibly because of the variety of industries coupled with the varying ways in which risk can present itself. The food industry will not have the same problems as a metal manufacturing industry. Even within the same industry, companies might not have the same problems. A seafood distributor will have different issues to assess than a fruit producer. These varying sources and types of risk have made it difficult to create guidelines in order to steer companies in the right direction to manage risk in their supply chain. Also, since there is a lack of general guidelines concerning risk management, every industry (even every company) has taken it upon themselves to develop their own guidelines.

The International Standard for Organization (ISO) seeks to remedy the varying levels of supply chain risk preparedness and quality by allowing companies to become universally certified in various ISO certifications to promote a unity of practice.

The steps for ISO certification are as follows:

Stage 1: Proposal stage—development of initial proposal of operational standards

Stage 2: Preparatory stage—preparing proposal for submission

Stage 3: Committee stage—committee review of proposal

Stage 4: Enquiry stage—questioning into company operations post-review

Stage 5: Approval stage—approval for ISO certification

Stage 6: Publication stage—ISO certification publicized and legitimized

Certifications, such as ISO 9001, and guidelines, such as ISO 31000, are best suited to prepare companies to handle risk (www.iso9001.com, 1) (www.iso.org, 1). Specifically, ISO 31000 (containing ISO Guide 73:2009 and ISO/IEC 31010:2009) seeks to establish guidelines for risk. ISO 31000 contains within it risk management principles, a framework for risk management, and a process for managing risk. It purports to be applicable to any organization regardless of size or sector and claims to increase the likelihood of improving the identification of opportunities and threats, thereby allowing an organization to more effectively allocate and use resources for risk treatment. ISO 31000 does not allow for certification but can aid with internal and external audits. These guidelines allow preparation for strategic, operational, and management risk and provide insight into the philosophy and practices that an organization might adopt in building an effective risk management system.

The ISO 9001 standards are also closely related to the approach used in quality improvement systems. In fact, the literature reports that there is a relationship between TQM and ISO 9001. TQM principles were incorporated into the ISO 9001 standard in the year 2000, so a significant relationship is likely to exist [24]. Studies have since shown a relationship between ISO 9001

and TQM. Psomas et al. [25] found that ISO 9001 certified companies achieve significant quality improvement through the implementation of the core process management practices characteristic of TQM. Sampaio et al. [26] found that most researchers in this area reason that ISO 9001 and TQM are very similar and should both be implemented in the organization together [27–29]. Others suggest that the structure of ISO 9001 can actually aid in the implementation of TQM practices [30]. Therefore, one would expect TQM principles and practices to be present in organizations that have achieved ISO 9001 certification. A few supply chain risk and quality studies do exist. Chapman et al. [31] studied delivery lead time variability and quality management. Tse et al. [32] explored quality and safety problems in the supply chain and introduced a supply chain risk management framework to reduce quality risk. Therefore, the relationship of supply chain risk management to quality management practices has not been widely studied. Based on this review of the literature, the following statements will be assessed empirically:

1. Organizations that are ISO 9001 certified have more mature risk management systems than organizations that are not ISO 9001 certified.
2. Organizations that have more mature risk management systems also demonstrate more mature quality management systems.

3. Data and research methodology

Data for this study came from a list of approximately 3000 United States–based organizations, compiled from the 2014 IAAR Directory, ISM-CV, and CSCMP member lists. These organizations cover varying locations throughout the United States spanning industries such as medical, manufacturing, government, etc. The approach used in this study is a survey questionnaire, conducted via e-mail, that includes demographic questions and questions specific to quality and risk management practices to gauge the relationship of quality management practices and risk management practices of organizations. The answers to the survey questions are then used to compare the level of risk management between organizations who more effectively use quality improvement practices and those who do not.

The survey was distributed to potential respondents using Qualtrics survey software. The survey components contained questions measuring the respondent's position within the organization, their familiarity with the organization's quality management and risk management practices, the organization's size, industry and position within the supply chain, where and how the organization manages risk, the company's quality management practices, and their ISO 9001 certification status. The answers to most of these questionnaire statements were then ranked with answers on a 1–5 scale (1 being the worst and 5 being the best). These responses were analyzed using SPSS software. Selected items from the survey questionnaire are provided in the Appendix. These statements include the number of responses in each category, as appropriate.

The measurement of quality management practice within organizations was taken from the work of Dellana and Kros [33]. They developed a set of 13 statements that cover the main

http://dx.doi.org/10.5772/65847

quality issues in an organization that could be found at any point in the supply chain and also in any type of industry. In their study, these 13 statements were split into two distinct constructs. Based on data from 565 respondents, 8 of the statements were assigned to the construct of internal-downstream quality (i.e., customer-related), while the remaining 5 statements were grouped under the construct of external-upstream quality (i.e., supplier-related). In their investigation, the internal-downstream construct was found to be associated with industry class and a measure of supply chain position, while the external-upstream construct was not. Therefore, in the current research, the measure of quality maturity is limited to the eight statements from Dellana's and Kros's previous work that made up the internal-downstream construct (ref. question 9 of the Appendix). The measurement of risk management practice in organizations was more challenging. No generally accepted measure of risk management maturity was found in the literature. Therefore, a set of statements was inspired and derived from a number of sources, including the ISO 31000 guidelines [34–42]. This resulted in the 15 statements listed in the questionnaire for question 7 of the Appendix.

4. Results

The results of the survey based on SPSS statistical software analysis are as follows. A total of 500 of the approximately 3000 potential respondents reached an active e-mail account. Of these 500, responses to the questionnaire were received from 40 individuals. Of these 40 responses, 18 respondents indicated their organization as ISO 9001 certified, while 20 responded that their organization was not ISO 9001 certified (two did not respond to the ISO certification question and therefore these two questionnaires were unusable). Analysis of nonresponse bias was performed by comparing the mean score for TQM of early versus late responders. There was no significant difference in the mean score between the two groups ($t = 1.40$, $p = 0.18$). However, given the small sample size, a comparison of means is by no means conclusive. Therefore, this study and its findings are considered preliminary and exploratory in nature. It should also be noted that this survey was conducted prior to the issuing of the ISO 9001:2015 standards, which incorporate a focus on risk management. The standards in force at the time of this research were ISO 9001:2008.

A general consideration of the survey responses related to demographics of the individual and organization follows. The job title and level of the respondents was generally varied. However, most of the respondents were at the manager level and above (79%). Most were somewhat to very familiar with their organization's supply chain risk management process (78%), with 38% reporting being very familiar. The overwhelming majority were somewhat to very familiar with their organization's quality management practices (94%), with 53% reporting being very familiar. The organizational size, based on number of employees, also varied quite a lot. Grouping into three categories of small, medium, and large yields relative percentages of 42, 29, and 29%, respectively. About 40% of respondent organizations have a department dedicated to supply chain risk management, while about 32% have each department responsible for assessing supply chain risk related to their particular function. Approximately, 35% are involved in the manufacturing part of the supply chain, while about 38% are involved in

distribution. The rest appear to be various services that relate to differing parts of the supply chain. The organizations fall mostly into the general industry categories related to manufacturing (34%) and services (34%). This is followed by transportation/distribution at about 21%. Finally, 47% of respondents reported their organizations are ISO 9001 certified while 53% reported not having ISO 9001 certification.

A factor analysis using principal component analysis was performed on the eight TQM-related questions to assess whether or not they were consistent with prior research. However, it should be noted that, for the purpose of factor analysis, the present sample did not meet the requirement that minimum sample size be at least 10 times the number of variables per Nunnally [43] ($n = 40$, $<10 \times 8 = 80$). That said, these 8 statements were already shown in Dellana and Kros's previous study to have met the requirements for factor analysis.

It was found that the eight TQM-related statements all loaded strongly on a single factor. Metrics related to factor analysis were run. The KMO measure of sampling adequacy (index = 0.863) and Bartlett's test of sphericity (p = 0.0000) were found to be acceptable. Therefore, the results presented herein were very consistent with prior work, which gave support for this research. The model met underlying assumptions (except sample size issues). The reliability of all eight questions together also was very good (Cronbach alpha reliability = 0.93). The TQM maturity variable value used in SPSS analysis was the average score for the eight statements out of a possible score of 5.0. Therefore, a higher average score indicates greater TQM maturity.

SCRM Factor 1 statements (Questionnaire statements i, k, l, m, n)	**SCRM Factor 2 statements** (Questionnaire statements b, d, e, f)
i. We follow the four step process of risk identification, analysis, education, and treatment	b. Risk management is an ad hoc process for us that occurs informally on an as-needed basis
k. We prioritize risk events based on severity of impact to our organization	d. Our risk management assessment is based on probability analysis of relevant risks
l. We involve our suppliers in identification and mitigation of potential supply chain risks	e. Our risk management assessment occurs at least annually
m. We encourage our suppliers to use a structured risk management process (e.g., ISO 31000)	f. Risk assessment is a quantitative process for us
n. We work with our customers to identify and mitigate potential supply chain risks	

Table 1. Results of factor analysis for supply chain risk management (SCRM) measurement.

Because the result was so positive for the TQM measure, factor analysis was also conducted on the supply chain risk management (SCRM) statements using principal components analysis in SPSS. Once again, the sample size requirement of Nunnally was not met in this case, so the analysis is preliminary. The KMO measure of sampling adequacy (index = 0.783) and Bartlett's test of sphericity (p = 0.0000) were found to be acceptable. Two factors were extracted using a varimax rotation and Kaiser normalization. The reliability of the statements in the two SCRM

http://dx.doi.org/10.5772/65847

factors held very well with Cronbach's alpha measures of 0.88 and 0.86 for Factor 1 and Factor 2, respectively. The two SCRM factor sets are described in **Table 1** and are synonymous with the survey statement designations. Factor 1 seems to describe more so "What" is managed during risk assessment, while Factor 2 seems to describe "How" this risk is assessed and corrected. An overall measure of SCRM combines these two sets of statements for a total of nine statements. In all cases of the risk maturity variables (i.e., SCRM, SCRM1 and SCRM2) the value used in SPSS analysis was the average score for the related statements out of a possible score of 5.0. Therefore, a higher average score indicates greater SCRM maturity.

4.1. Research question 1

The first research question explores whether organizations that are ISO 9001 certified have more mature risk management systems than organizations that are not ISO 9001 certified.

The relationship of ISO certification with SCRM was measured with the dependent variable entered as the respective risk management variables (i.e., SCRM1, SCRM2, and SCRM) using a univariate linear regression in SPSS with ISO 9001 certification as an independent variable. In all cases, the number of employees, supply chain position, and TQM score were also entered as control variables. Because of low frequencies by category, some data consolidation was performed. The number of employees was collapsed into small, medium, and large groups. Supply chain position was generally sorted into manufacturing and distribution and other.

The results of the regression analysis are given in **Table 2**. Backward removal of variables in SPSS resulted in the ISO 9001 reaching a significance of at best $p = 0.064$. This was for the case of the SCRM overall. SCRM2 was second best with a significance of $p = 0.079$. Neither of these make the generally accepted threshold of 0.05 significance level and were, therefore, not included in the final SPSS models. The models were also run with all control variables excluded to determine whether this would make a difference to the result. The outcome was similar to the model that included control variables, with SCRM1 at a significance of $p = 0.806$, SCRM2 at a significance of $p = 0.061$, and SCRM at a significance $p = 0.248$. Therefore, the evidence is weak for a conclusion that ISO 9001 certified organizations have more mature risk management systems than non-ISO 9001 certified organizations.

A breakdown by industry type had been a goal of this research, but the sample size was too small to accommodate a rigorous statistical analysis. However, it is worth noting that the split of ISO and non-ISO certified organizations was quite different between organizations classified as manufacturing (9 certified, 4 noncertified) and those classified as service (2 certified and 11 noncertified). Distribution-related organizations were more evenly split (five certified and three noncertified). **Table 3** shows a breakdown of the scores for SCRM and TQM by major industry type. The scores for SCRM seemed to favor the service organization over manufacturing, especially for SCRM2. Although ISO 9001 does not clearly differentiate regarding SCRM maturity, there is at least the implication in this research that industry type may be a differentiator. A similar analysis could be conducted by supply chain position. It might be expected that organizations further downstream in the supply chain would exhibit greater SCRM maturity than those upstream given the lengthening of the supply chain heading

downstream, which increases the chance of risk event occurrence that disrupts the supply chain.

Variable	Beta	t	Significance
SCRM1	0.209	1.519	0.139
• SC position	0.630	4.516	**0.000**
• **TQM**	0.117	0.743	0.464
• ISO 9001	−0.070	−0.467	0.644
• Number employees			
SCRM2	0.117	0.625	0.537
• SC position	0.280	1.653	0.109
• TQM	0.310	1.814	0.079
• ISO 9001	−0.059	−0.324	0.749
• Number employees			
SCRM	0.154	0.903	0.374
• SC position	0.547	3.469	**0.002**
• **TQM**	0.304	1.928	0.064
• ISO 9001	−0.119	−0.714	0.481
• Number employees			

Table 2. Linear regression analysis results for SCRM variables versus ISO 9001.

Score category	Manufacturing (n = 13)	Service (n = 13)	Distribution (n = 8)
SCRM	3.13	**3.71**	3.56
SCRM1	3.32	**3.60**	3.58
SCRM2	2.84	**3.84**	3.53
TQM	4.02	4.05	**4.32**

Table 3. Mean SCRM and TQM scores by general industry type.

4.2. Research question 2

The second research question explores whether organizations that have more mature supply chain risk management systems also demonstrate more mature quality management systems. The simple linear regression analysis of **Table 2** incorporates the TQM score as an independent variable in the analysis of relationships with SCRM variables. TQM was found to be strongly significantly related to SCRM1 ($p < 0.000$) and SCRM overall ($p = 0.002$). Both of these met the

http://dx.doi.org/10.5772/65847

generally accepted threshold of 0.05 significance level and were, therefore, included in the final SPSS models. SCRM Factor 1 statements tend to describe the actual risk management process details, outlining the process steps and who is involved. SCRM Factor 2 statements are more general and have more to do with the timing and nature of the process. The models were also run for SCRM variables with TQM as the only independent variable in each case to determine whether this would make a difference to the result. The outcome was similar to the model that included control variables with SCRM1 at a significance of $p < 0.000$, SCRM2 at a significance of $p = 0.248$, and SCRM at a significance $p = 0.003$. This suggests that the presence of a mature TQM program may signal a more mature supply chain risk management system based on the actual process scope and steps.

Analysis was also run on the correlation between TQM and SCRM1 specifically to see where there were strong relationships between specific quality management statements and supply chain risk management maturity statements. SCRM1 statements i, k, and l were all positively correlated to TQM statements a, b, c, d, e, g, and h (see **Table 1** for the specific questions). SCRM1 statement m was positively correlated to TQM statements a, b, c, g, and h. SCRM1 statement n, b, g, h was positively correlated to TQM statements b, g, and h. Therefore, statement f, "Quality metrics or standards are kept," was not correlated to any SCRM1 statements. This analysis suggests that TQM is strongly related to most of the SCRM1 statements, with only a few weak relationships. Statement 6 stands out as unrelated to SCRM1, while statement n of SCRM1, "We work with our customers to identify and mitigate potential supply chain risks", stands out as only weakly related to TQM.

5. Conclusions

These results suggest that ISO 9001 is not strongly related to supply chain risk management maturity. There was not a strong relationship between ISO certification and company's preparedness level of SCRM and TQM. However, TQM is clearly related to SCRM1, no matter the variables that are included. The more mature a company's TQM practices, the more mature the company is in terms of SCRM as identified by the SCRM Factor 1 (which describes the actual risk management process details, outlining the process steps and who is involved). Most of the statements comprising this factor were positively correlated to the TQM statements (denoting total quality management practices). In particular, these included the following SCRM1 statements:

- (i) "We follow the four step process of risk identification, analysis, education, and treatment;"
- (k) "We prioritize risk events based on severity of impact to our organization;"
- (l) "We involve our suppliers in identification and mitigation of potential supply chain risks;"
- (m) "We encourage our suppliers to use a structured risk management process (e.g., ISO 31000)."

The statement related to customers (i.e., "We work with our customers to identify and mitigate potential supply chain risks") was only weakly associated to TQM.

The lack of significance between ISO9001:2008 and the two factors of SCRM, based on survey results and analysis, suggests that the standards for ISO 9001:2008 and 31000 do not provide a significant advantage to risk assessment and management.

Though they provide positive frameworks for overall company structure and project development, ISO 9001:2008 was not equipped to provide a framework for risk management that eclipses alternative methods. It should be noted that ISO 31000 is a new addition and has not yet been fully developed. Before the implementation of ISO 31000, risk was not explicitly addressed under ISO standards.

6. Limitations and further research

The number of respondents in this study was relatively small and necessarily reduces the power of the statistical tests in this study. The low response rate also introduces the potential for nonresponse bias. Therefore, the conclusions should be treated with caution. However, the results suggest that a higher response rate would indicate a similar pattern of results based on prior studies, in particular related to the TQM maturity metric.

To further aid such a study, case studies may be performed to monitor the daily processes of companies who are and are not ISO certified. This would provide more of an in depth analysis of the effects of ISO certification on risk management. E-mail provided certain limitations to accessibility, because of company filtering software and lack of personal interaction. Joint participation with the ISO organization into the effectiveness of their certifications would also lead to a better understanding and greater ease of access.

However, the results of this initial survey and analysis suggest that ISO does not play a significant role in risk assessment and correction. Though, the more mature a company is able to be in terms of their total quality management procedures, the better equipped it will be to handle supply chain risk.

Further study should be conducted to seek a larger sample size in order to assess the reliability of the results in this study. It would also help shed light on differences by industry type and supply chain position. However, this will necessarily involve the newer standards for ISO 9001:2015. Therefore, it is unlikely that this study can be replicated. It would, nonetheless, still be of interest to assess the degree to which the new standards have impacted on organizational effectiveness in managing risk in the supply chain. Since the new standards now focus attention specifically on risk management, it should signal to customers that ISO 9001:2015 certification is a good supplier prequalifier when it comes to supply chain risk management and not just quality. Or at least that would be the expectation. Further study is needed to confirm this.

http://dx.doi.org/10.5772/65847

Appendix: Selected survey questions (number responding given in parentheses)

1. What is your current job title/level?

- CEO/President (7)
- Vice President (8)
- Director (4)
- Manager (11)
- Experienced Staff Member (3 or more years of experience) (3)
- Entry Level Staff Member (less than 3 years of experience) (1)
- Other. Specify: (4)

2. About how many employees does your organization have?

- 20 or fewer (5)
- 21 to 100 (11)
- 101 to 500 (6)
- 501 to 5000 (5)
- 5001 to 10,000 (3)
- More than 10,000 (8)
- Not sure

3. Which industry category does your organization best fit into?

a. Aerospace manufacturing (2)

b. Automotive manufacturing (1)

c. Chemicals/plastics manufacturing (2)

d. Computer equipment manufacturing (0)

e. Electronics/electrical manufacturing (1)

f. Fabricated metal products manufacturing (2)

g. Food products manufacturing (0)

h. Government/public administration (1)

i. Machinery manufacturing (Industrial/Commercial) (4)

j. Pharmaceuticals manufacturing (1)

k. Retail trade/merchandising (0)

l. Services: Educational services (1)

m. Services: Professional/scientific/technical (12)

n. Transportation (1)

o. Utilities (1)

p. Warehousing/distribution (7)

q. Wholesale trade (2)

r. Other. Specify: (1)

4. How does your organization conduct supply chain risk analysis? (select one)

- We outsource our supply chain risk analysis to a consulting firm. (1)
- We have a dedicated Risk Management Department (i.e., an internal consultant) that performs supply chain risk analysis for the organization. (9)
- Each department is responsible for assessing supply chain risk related to their particular function. (12)
- We have a specific department (e.g., procurement) that is charged with managing supply chain risk. (Please specify department:) (6)
- Other. Specify: (9)

5. Which of the following positions best describes your organization's business function in the supply chain?

- Raw/basic material manufacture (1)
- Component/sub-assembly manufacture (1)
- Final product manufacture (11)
- Warehousing/distribution (14)
- Retail (0)
- Other. Specify: (10)

6. How familiar are you with your organization's supply chain risk management process?

- Very unfamiliar (7)
- Unfamiliar (1)
- Somewhat familiar (7)
- Familiar (8)
- Very familiar (14)

http://dx.doi.org/10.5772/65847

7. The following statements relate to your organization's supply chain risk management system. Please indicate the degree to which you agree or disagree with each statement in relation to your organization.

(1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree)

a. Risk management is built into our planning process.

b. Risk management is an ad hoc process for us that occurs informally on an as-needed basis.*

c. Our risk management assessment occurs on a regular schedule.

d. Our risk management assessment is based on probability analysis of relevant risks.

e. Our risk management assessment occurs at least annually.

f. Risk assessment is a quantitative process for us.

g. Our risk management process is assessed for continual improvement.

h. Top level management is involved in our risk management process.

i. We follow the four step process of risk identification, analysis, education, and treatment.

j. We prioritize risk events for treatment based on results of risk analysis.

k. We prioritize risk events based on severity of impact to our organization.

l. We involve our suppliers in identification and mitigation of potential supply chain risks.

m. We encourage our suppliers to use a structured risk management process (e.g., ISO 31000).

n. We work with our customers to identify and mitigate potential supply chain risks.

o. We have a process in place for corrective feedback to our risk management process.

*Coded in reverse in the analysis.

8. How familiar are you with your organization's quality management practices?

- Very unfamiliar (1)
- Unfamiliar (1)
- Somewhat familiar (5)
- Familiar (9)
- Very familiar (18)

9. The following statements relate to your organization's quality management system. Please indicate the degree

to which you agree or disagree with each statement in relation to your organization.

(1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, 5 = Strongly agree)

- **a.** Customer satisfaction/service data for key indicators are actively accumulated and analyzed.
- **b.** Customer data are used to make internal quality improvements.
- **c.** Customer complaints are tracked and managed through a formal complaint system with feedback.
- **d.** Customer service standards are kept.
- **e.** Employee use of quality-related problem solving techniques is actively supported (ex. DMAIC, control charting, etc.).
- **f.** Quality metrics or standards are kept.
- **g.** Team-based quality improvement is actively supported.
- **h.** Top management has a demonstrated commitment to quality.

10. Is your organization certified for ISO 9001 or equivalent?

- Yes (18)
- No (20)

Acknowledgements

This chapter is built on prior research from "The Relationship of Organizational Quality Management Practice with Supply Chain Risk Management Program Maturity", Thesis for East Carolina University Honors College

Author details

Tyler Florio

Address all correspondence to: tylerwflorio@gmail.com

College of Business, East Carolina University, Greenville, USA

http://dx.doi.org/10.5772/65847

References

[1] Punniyamoorthy M., Thamaraiselvan N., Manikandan L., Assessment of supply chain risk: scale development and validation. Benchmarking: An International Journal. 2013; 20(1):79–105.

[2] Wildgoose N., Brennan P., Thompson S., Understanding your supply chain to reduce the risk of supply chain disruption. Journal of Business Continuity & Emergency Planning. 2012; 6(1): 55–67.

[3] Van Miegham J., Risk management and operational hedging: an overview. In: Kouvelis P. et al. Editors. Handbook of integrated risk management in global supply chains. Hoboken, NJ: John Wiley & Sons; 2011. p. 13–50.

[4] Chen J., Sohal A., Prajogo D., Supply chain operational risk mitigation: a collaborative approach. International Journal of Production Research. 2013; 51(7): 2186–2199.

[5] Gualandris J., Kalchschmidt M., Product and process modularity: improving flexibility and reducing supplier risk failure. International Journal of Production Research. 2013; 51(19):5757–5770.

[6] Scholten K., Scott P.S., Fynes B., Mitigation processes—antecedents for building supply chain resilience. Supply Chain Management: An International Journal. 2014; 10(5/6): 211–228.

[7] Golgeci I., Ponomarov S.Y., Does firm innovativeness enable effective responses to supply chain disruptions? An empirical study. Supply Chain Management: An International Journal. 2013; 18(6):604–617.

[8] Kunimatsu L., Risk Management Basics—ISO 31000 Standards. 2013. p. 1–17 [cited 2016 Aug 30]. Available from: http://www.secureworldexpo.com/2011/detroit/Louis_Kunimatsu.pdf.

[9] Fowles J., Handbook of futures research. Connecticut: Greenwood Press; 1978.

[10] Moeller R.R., COSO enterprise risk management: establishing effective governance, risk, and compliance processes. Hoboken, NJ: Wiley; 2011.

[11] Cavanaugh, T. Preparedness in the Private Sector. The Conference Board. December 2008;Report Number R-1436-08-RR. Available from: https://www.conference-board.org/publications/publicationdetail.cfm?publicationid=1592.

[12] Bhatia G., Lane C., Wain A., Building resilience in supply chains: an initiative of the risk response network in collaboration with Accenture. World Economic Forum: Cologny/Geneva Switzerland; 2013.

[13] Ladbury A., Supply-chain risks misunderstood, mismanaged: report. Business Insurance. 2008; 42(2):23.

[14] Chacon N., Doherty S., Hayashi C., Green R., New models for addressing supply chain and transport risk: an initiative of the risk response network in collaboration with Accenture. World Economic Forum: Cologny/Geneva Switzerland; 2012.

[15] Ghadge A., Dani S., Kalawsky R., Supply chain risk management: present and future scope. International Journal of Logistics Management. 2012; 23(3):313–339.

[16] Kumar S. Managing risks in a relief supply chain in the wake of an adverse event. OR Insight. 2011; 24(2):131–157.

[17] Aguiar Y.M., Assessing the CARVER+S risk management model of terrorism preparedness in business continuity planning. Northcentral University, USA: PhD dissertation; 2011.

[18] Skipper J.B., Hanna J.B., Gibson B.J., Alabama power response to Katrina: managing a severe service supply chain disruption. Arden: Jordan Whitney Enterprises Inc; 2010.

[19] Ellegaard C., Supply risk management in a small company perspective. Supply Chain Management. 2008; 13(6):425–434.

[20] Kern D., Moser R., Hartmann E., Moder M., Supply risk management: model development and empirical analysis. International Journal of Physical Distribution & Logistics Management. 2012; 42(1):60–82.

[21] Pettit T.J., Supply chain resilience: development of a conceptual framework, an assessment tool and an implementation process. The Ohio State University, USA: PhD dissertation; 2008.

[22] Finch P., Supply chain risk management. Supply Chain Management. 2004; 9(2):183–196.

[23] Zhao L., Huo B., Sun L., Zhao X., The impact of supply chain risk on supply chain integration and company performance. Supply Chain Management: An International Journal. 2013; 18(2):115–131.

[24] Fotopoulos C., Psomas E., The use of quality management tools and techniques in ISO 9001:2000 certified companies: the Greek case. International Journal of Productivity and Performance Management. 2009; 58(6):564–580.

[25] Psomas E.L., Fotopoulos C.V., Kafetzopoulos D.P., Core process management practices, quality tools and quality improvement in ISO 9001 certified manufacturing companies. Business Process Management Journal. 2011; 17(3):437–460.

[26] Sampaio P., Saraiva P., Rodrigues A.G., ISO 9001 certification research: questions, answers and approaches. International Journal of Quality & Reliability Management. 2009; 26(1):38–58.

http://dx.doi.org/10.5772/65847

[27] Escanciano C., Fernandez E., Vasquez C., Influence of ISO 9000 certification on the progress of Spanish industry towards TQM. International Journal of Quality & Reliability Management. 2011; 18(5):481–494.

[28] Gotzamani K., Tsiotras G., An empirical study of the ISO 9000 standards' contribution towards total quality management. International Journal of Operations & Production Management. 2001; 21(10):1326–1342.

[29] Dwyer G., Business excellence versus ISO 9000 in an Irish context-which delivers? Managerial Auditing Journal. 2002; 17(7):404–411.

[30] Magd H., Curry A., ISO 9000 and TQM: are they complementary or contradictory to each other. The TQM Magazine. 2003; 15(4):244–256.

[31] Chapman P., Bernon M., Haggett P., Applying selected quality management techniques to diagnose delivery time variability. International Journal of Quality & Reliability Management. 2011; 28(9):1019–1040.

[32] Tse Y.K., Tan K.H., Chung S.H., Lim M.K., Quality risk in global supply network. Journal of Manufacturing Technology Management. 2011; 22(8):1002–1013.

[33] Dellana S.A., Kros J.F., An exploration of quality management practices, perceptions and program maturity in the supply chain. International Journal of Operations and Production Management. 2014; 34(6):786–806.

[34] Gray C.F., Larson E.W., Project management: the managerial process. 4th ed. New York: McGraw-Hill Companies; 2008.

[35] Foster S.T., Managing quality: integrating the supply chain. 5th ed. New Jersey: Pearson Education Inc.; 2013.

[36] Kouvelis P., Dong L., Boyabati O., Li R., Integrated risk management: a conceptual framework with research overview and applications practice. In: Kouvelis P. et al. Editors. Handbook of integrated risk management in global supply chains. Hoboken, NJ: John Wiley & Sons; 2011. p. 3–12.

[37] ISO 31000:2009. Risk management—principles and guidelines. ISO; 2009 [cited 2016 Aug 21]. Available from: https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-1:v1:en.

[38] Leitch M., ISO 31000:2009—The new international standard on risk management. Risk Analysis: An International Journal. 2010; 30(6):887–92.

[39] DeRosier J., Stalhandske E., Bagian J.P., Nudell T., Using health care failure mode and effect analysis: The VA National Center for Patient Safety's Prospective Risk Analysis System. The Joint Commission; 2002, p. 248–67.

[40] SCRLC, Supply chain risk management: a compilation of best practices. Supply Chain Risk Leadership Council; 2011.

[41] Thompson A. Risk Management Framework. Government of South Australia, Adelaide: Department of Communities and Social Inclusion; 2009.

[42] COSO., Strengthening enterprise risk management for strategic advantage. Committee of Sponsoring Organizations of the Treadway Commission; 2009 [cited 2016 Aug 21]. Available from: www.coso.org.

[43] Nunnally J.C., Psychometric theory. 2nd ed. New York: McGraw-Hill; 1978.

Chapter 7

ALAMEDA Ecosystem: Centering Efforts in Software Testing Development

José González Enríquez,
Julián Alberto García-García,
Francisco José Domínguez-Mayo and
María José Escalona Cuaresma

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66043

Abstract

One of the most important and critical aspects to improve the quality assurance in software is to improve the testing process by utilizing techniques and tools, which will enhance the software testing process, making it more effective and efficient. This chapter presents ALAMEDA ecosystem, a software package that centers its efforts in software testing development and is a result from a real-world project. ALAMEDA provides support to lifecycles focused on the generation, implementation, and testing organization from the earliest stages of software development. In addition, the ecosystem provides an environment of rating the degree of compliance of organizations with the International Standard for Testing ISO/IEC-29119. It is proposed as a tool to use during the various iterations that may occur in an agile software development process.

Keywords: testing development, model-driven testing, quality assurance

1. Introduction

The area of Information Technology and Communications (ICT) is an essential element of innovation in other business sectors such as industrial, logistics, health, etc., where the effectiveness and efficiency in software development is critical for a fundamental and safe operation.

The process of software development involves a series of activities in which the chances of a human error is high (mistakes can happen in the beginning of the process, in which the objectives can be inadequately specified, as well as during later steps). Therefore, it is important to research and define techniques and methods (as well as supporting tools) in order to detect and solve any area of special attention as soon as possible. In this sense, over recent decades, a number of studies have analyzed the cost of resolving software defects. The cost increases with the time of the defect in the system [1]. As such, the cost of rectifying an error is lower, the earlier the problem is detected.

The budget of software projects is critical and limited in many cases [2]. For this reason, both research institutions and enterprises look for solutions to improve quality assurance of the software development process [3]. The finding of methods and tools to improve the quality, reduce costs and increase the guarantee of results becomes an essential aim for enterprises and development teams [4].

One of the most important and critical aspects to improve the quality assurance in software is to improve the testing process by utilizing techniques and tools, which will enhance the software-testing process, making it more robust. It is a systematic method to determine inter-related factors and parameters affecting a process and its desired output [5]. It is usually performed for one of two reasons: to detect defects and to estimate reliability. The key to software testing is trying to find the modes of failure—something that requires testing the code on all possible inputs and depending on the environment it can be implemented at any time during the development process [6].

This chapter presents a testing solution (named ALAMEDA) to innovate in all these aspects related to software testing. ALAMEDA offers to project managers and developers a systematic and tool-based approach for managing (i.e., defining, controlling, measure, etc.) test cases. This systematic and automatic management allows the detection and correction of software errors before they become costly interruptions in the production line.

The ALAMEDA ecosystem aims to provide a set of applications which allows software development companies to ensure the management, automation, and integration of test cases from an early stage of the develop lifecycle. For this purpose, ALAMEDA integrates the well-known Model-Driven Engineering (MDE) [7, 8] paradigm on software testing and its automation of software and testing development by use of the NDT (Navigational Development Techniques) methodology [9] and NDT-Suite [10].

MDE has been used as a direction vector of the ALAMEDA ecosystem because it is one of most entrenched paradigms within the software engineering area. As such, on the one hand it will help to define the approach and on the other hand, it will manage the conceptual complexity the process engineering area entails. In addition, MDE has been also successfully applied to real software environments and it has shown a considerable area of impact on reducing time-to-market and improving software product quality. For instance, in the business process management [11], healthcare, software product lines [12], or web engineering among others.

http://dx.doi.org/10.5772/66043

The remainder of this chapter is organized as follows: Section 2 presents the theoretical foundations of ALAMEDA. Section 3 presents the ALAMEDA ecosystem for software testing. Finally, Sections 4 and 5 briefly discuss some related work and highlight conclusions.

2. Theoretical foundations of ALAMEDA

The model-driven software development has become one of the most important paradigm in computer engineering. Indeed, it has shown a considerable impact on reducing time-to-market and has improved overall quality.

However, the development of high-quality software solutions does not only require systematic processes of software development, but it also requires systematic processes of software testing. At present, this latter aspect has not been attracting attention within the software industry.

It is a fact that if software companies do no invest resources and efforts in the process of software testing, it is quite likely that there will be a waste of money in the short term because of a large volume of errors. Therefore, if a systematic process is not used in tasks of testing, software developers will spend most of the time to check quality and application functionality rather than developing new projects [13].

The ALAMEDA ecosystem aims to innovate in the systematic and automatic application of the process of software testing in the real environment. For this purpose, the ecosystem presented herein takes into account the new standard for software testing (ISO/IEC/IEEE 29119) and a well-validated methodology named NDT. The latter proposes techniques and mechanisms to generate test cases from the early stages of the software lifecycle. These theoretical foundations are explained in the next sections.

2.1. ISO/IEC/IEEE 29119 software testing

The International Software Testing Standard (ISO/IEC/IEEE 29119) [14] is a standard created between 2013 and 2015. It was created in order to unify all knowledge on software testing and establish basic knowledge concerning this discipline.

On the one hand, this standard provides useful methodological and conceptual guidelines from the perspective of professional or software providers. In addition, it provides (offers) a common terminology, certifications, well-defined techniques, professional qualification, continuous improvement of the process of software testing, inter-operability, and consistency, among others.

On the other hand, and from the perspective of customers, ISO/IEC/IEEE 29119 provides customer confidence on their software provider, an industry benchmark for good practice, and a contractual relationship between the customer and the software provider, among others.

ISO/IEC/IEEE 29119 focuses on a process model of three levels based on the risk factor in the software-testing phase. This model provides guidelines on strategies and policies so as to

manage the process of software testing (i.e., defining testing plans, monitoring and controlling software tests, establishing the test environment, executing tests, etc.). All these aspects have been described in detail in five volumes, as follows:

- ISO/IEC 29119-1: Concepts & Definitions. It introduces the vocabulary on which all standards in the 29119 series are built. It also provides examples of the application of each concept in practice.
- ISO/IEC 29119-2: Test Processes. It defines a generic process model for software testing that can be used to govern, manage, and implement software testing in any organisation, project, or testing activity. This model is based on three layers: (i) organizational test specifications (e.g., organizational test policy, organizational test strategy); (ii) test management; and (iii) dynamic testing.
- ISO/IEC 29119-3: Test Documentation. It defines templates for test documentation that cover the entire software testing life cycle.
- ISO/IEC 29119-4: Test Techniques. It defines techniques and methods to design test cases.
- ISO/IEC 29119-5: Keyword Driven Testing. It defines guidelines for supporting keyword-driven testing which is a way of describing test cases by using a pre-defined set of keywords. These keywords are names which are associated with a set of actions that are required to perform a specific step in a test case. By using keywords to describe test steps instead of natural language, test cases can be made easier to understand, to maintain, and to automate.

2.2. Model-driven engineering on software testing

The model-driven engineering paradigm [15] came up in order to tackle the complexity of platforms and the inability of third generation languages to relieve this complexity. It effectively expresses the domain concepts of the problem. This new paradigm, apart from raising the level of abstraction [16], intends to increase automation during the life cycle of software development.

MDE works, as the primary form of expression, with definitions of models (which can be defined, e.g., by means of Unified Modeling Language (UML) [17]) and transformations (which can be defined, e.g., by means of Query/View/Transformation (QVT) [18]) among these models which entail the production of other models. Both elements are essentials to apply MDE. In addition, it is important to take into account the concepts of metamodel and transformation rules. On the one hand, a metamodel is composed of a set of basic elements, its relationships, and its semantic constraints to build well-defined models. On the other hand, a transformation between two models defines a set of relationships between elements of metamodel A (source) and elements of metamodel B (target) [19].

Although MDE offers a theoretical framework, it has not been standardized exactly in these terms in order to facilitate the application of MDE in real projects. OMG presents MDA, which stands for model-driven architecture [20], as a platform to support the MDE paradigm. MDA proposes to base the software development on models which make transformations in order to generate a code or another model with characteristics of a particular technology (or lowest

http://dx.doi.org/10.5772/66043

level of abstraction). As transformations go on, it may be noted that the models become more concrete and the abstract model changes into another one compatible with a particular technology or platform. MDA is based on four types of levels or models. These models are (of highest to lowest level of abstraction):

- The CIM (computation-independent model) level is considered the highest level of business model and is associated with the most abstract level. It focuses on requirement specification and regards that anyone who knows the business and its processes can understand a CIM model, as this avoids any contact with a specific system.
- The PIM (platform-independent model) level represents the business process model and system structure, without any reference to the platform on which the application will be implemented. It is usually the entry point for all the support tools for MDA.
- The PSM (platform-specific model) level explicitly relates to the platform on which the system will be implemented, for example, with operating systems, programming languages, or middleware platforms, among others.
- The Code level refers to the codification and suitable implementation of the system.

MDE in general (or MDA in particular) has been successfully applied in different environments and research areas [21], software testing is one of them. In fact, this particular case is named as model-driven testing (MDT).

MDT is becoming more standardized in order to approach the automation of software testing, thanks to different research works and proposals [22–24]. This approach can significantly reduce the most painstaking cycle of all software development efforts—testing. Testing currently comprises between 30 and 70% of all software development projects, which calls for a significant amount [25] of resources that must be properly managed. Therefore, MDT and supporting tools enable software developers and testers to become far more productive and reduce the time-to-market, while maintaining high standards of software quality.

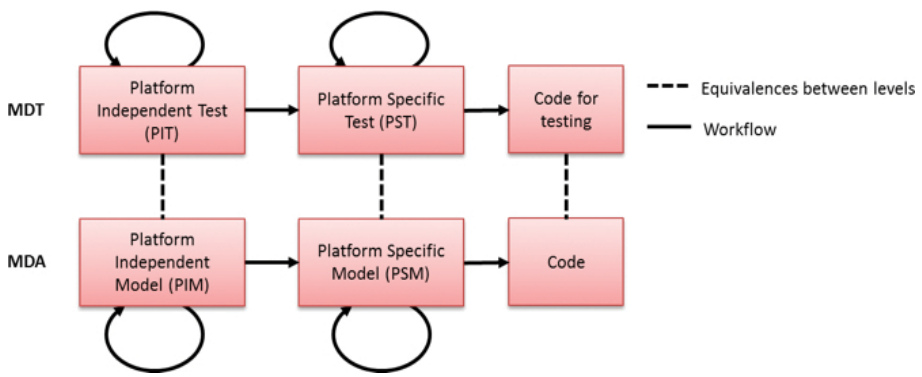

Figure 1. General scheme for generation of test cases integrated in ALAMEDA.

Figure 1 shows the vision of MDT from the perspective of MDA. As this suggests, models belonging to the PIM level can be transformed into models belonging to the PSM level, getting the code from the systems derived from PSM through another transformation step. This is the general philosophy of work proposed by MDA, but if this one is applied to software testing, some concepts change. In this case, the process of testing starts with models of System's Test which belong to the PIM level; or also named platform independent test (PIT). The quality team will elaborate on these tests to provide additional information. Next the PSM, which contains specific information about the implementation of the system, can be translated into the Platform Specific Test (PST), using the previous information from PIT and the knowledge of the quality team. Finally, these tests are morphed into codes for testing the system, ensuring the quality of the test and covering all the project-specific requirements.

All these theoretical foundations of MDT have been systematically applied and adapted within the ALAMEDA ecosystem. For this purpose, ALAMEDA implements a methodological framework to generate case tests from functional requirements (FR) as well as managing and controlling these case tests. This framework is based in NDT, a well-contrasted methodology. The foundations of NDT and how this one has been integrated into ALAMEDA are detailed in Section 2.3.

2.3. The NDT methodology from the perspective of software testing

NDT is a web methodology and is covered by the MDE paradigm. At present, NDT supports classical and agile software development lifecycles. For this purpose, NDT defines a set of metamodels for each phase of the software lifecycle (viability study, requirements, analysis, design, implementation, testing and maintenance) and it establishes a set of transformation rules (based on QVT) to generate models from other ones systematically. This implies a lower cost for the software development because of the automation of the process.

The application of MDE-based methodologies (such as NDT) and particularly, the application of transformations among models may become monotonous and very expensive if there are no supporting tools. These tools automate the process in order to get all the potential of MDE which provides a practical and useful environment to the enterprises. This aspect is one of the virtues of NDT by which this methodology has been applied successfully in many real projects. In this sense, NDT defines a set of supporting tools grouped in NDT-Suite [26]. The main tools in NDT-Suite are: (i) NDT-Profile, which defines UML profiles for each NDT metamodel; (ii) NDT-Quality [27], which measures the quality of use of NDT and checks semantic constraints of NDT; and (iii) NDT-driver [28], which allows running each QVT transformations between models in an automatic manner (for instance, this tool generates test cases from the specification of FRs of a project).

From the perspective of the process of software testing, ALAMEDA uses mechanisms defined by NDT to generate test cases from functional requirements defined in the first phase of the software lifecycle (i.e., the requirement phase). ALAMEDA also proposes how functional testing can be deeply improved by means of *early testing*. This paper aims to briefly present how this mechanism of generation of test cases works [29, 30].

http://dx.doi.org/10.5772/66043

As mentioned above, ALAMEDA implements the concept of *early testing* which means that test cases can be generated from early stages of the software lifecycle. It does not need to define test cases when the system is developed (classic software development lifecycle). In this sense, the generation of test cases of ALAMEDA is based on a four-step process (**Figure 2**):

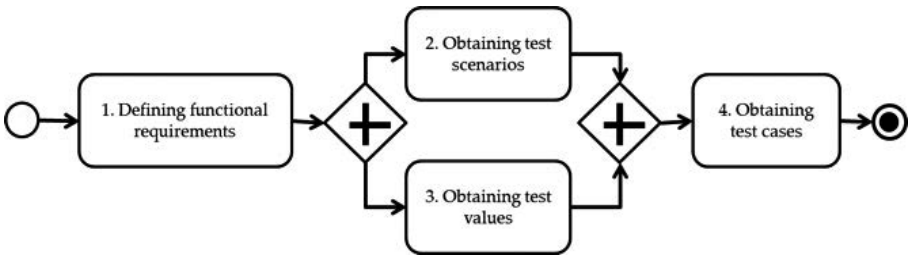

Figure 2. Test generation of ALAMEDA ecosystem.

1. Defining functional requirements. In this step, the user defines the functional requirements of his/her system. For this purpose, the user should use the concrete syntax of the FR metamodel of NDT. This definition is performed using NDT-profile (which has been integrated in ALAMEDA).

2. Obtaining test scenarios. In this stage, ALAMEDA manages each FR as a graph or a state machine. Taking into account this consideration, the methodology applies classic algorithm of path finding in a state machine in order to generate paths. Each path will be a scenario designed together with the system. At the same time, each scenario is a potential test case for assessing the right implementation and functionality thereof.

3. Obtaining test values. This step of the process consists of applying the category-partition method [31] to FRs. This method is based on identifying categories and partitions and to then generate combinations among such partitions. In the context of FRs, a category is any point for which the FRs defines an alternative behavior. Once all categories and partitions are identified, a combination between them becomes a potential test case.

4. Obtaining test cases. Once generated each scenario and possible values, ALAMEDA combines both results to generate final test cases. This step and the previous are systematically generated thanks to the QVT transformation rules defined by NDT. In addition, these transformations are performed using NDT-profile (which has been integrated in ALAMEDA).

3. ALAMEDA ecosystem

The ALAMEDA ecosystem for software testing consists of:

- a methodological framework based on the ISO 29119 International Testing Standard and

- a technological support that allows carrying out of the above methodological framework and is composed of a web platform and a desktop platform. It is worth mentioning at this point that both allow the application of MDT and the automation of development tests by the use of NDT-Suite, respectively. In addition to this, they incorporate a set of tools to automate the dynamic tests and perform the management and quality of their software in an integrated way.

Concretely, the ALAMEDA ecosystem offers a platform as a service (PaaS) ready-to-use, a quick and a progressive methodological implementation and adaptation. It is also flexible and easily scalable and offers an alignment with the ISO 29119 International Testing Standard.

3.1. Methodological framework

In this section, the different actors and activities involved in the ecosystem are shown. The different cases performed by an actor are also described.

Figure 3 shows the three main types of actors who inherit different subtypes.

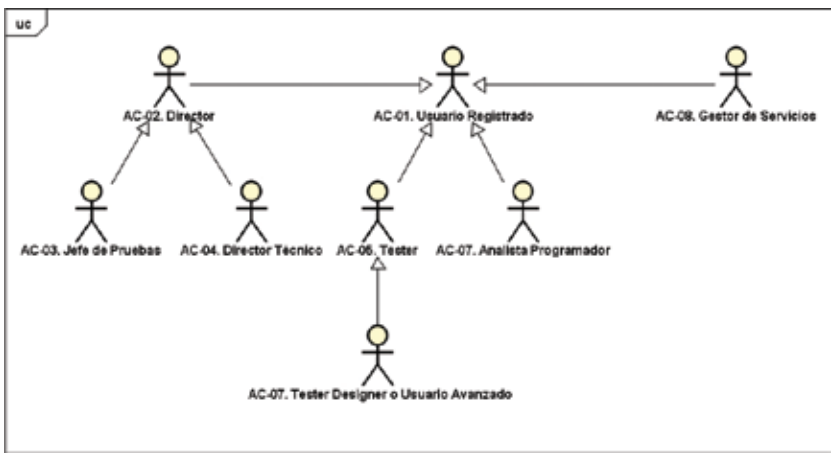

Figure 3. Actors of ALAMEDA ecosystem.

- **AC-01. Registered user**: Actor who represents a team member and is also a registered user of ALAMEDA technological tools. This actor, is nested in:
  - **AC-02. Director**: Actor who represents a registered user of director or head of the corporation type to ensure product quality software. This actor is divided into two subtypes:
    - **AC-03. Test manager**. Actor who must manage a testing a project and initiating the necessary meetings to ensure its success.
    - **AC-04. Technical manager**. Actor who represents the person with overall responsibility for the management or conduct of the projects concerned.

http://dx.doi.org/10.5772/66043

- **AC-05. Tester**. Actor representing the user or users responsible for testing the system. In turn, this actor can be defined as a specific type:
  - **AC-06. Tester designer or advanced user**. Actor who represents an advanced user tester and is responsible for designing the tests to be performed by the system.
- **AC-07. Analyst-programmer**. Actor who represents a team member of the software development project.
- **AC-08. Services manager**. Actor who represents an administrator or a manager user.

There are 13 functionalities that define the methodological framework of ALAMEDA ecosystem; those are:

- **FR-01. Manage methodological guide**. In this scenario, two types of actors are involved: the technical manager and all the other actors who are involved in ALAMEDA. The first step to be undertaken by the technical manager is a self-assessment regarding compliance with ISO 29119. Next, it is necessary to determine whether to make improvements in the methodology of the organization. If this is not necessary, the self-assessment of compliance with the standard will be performed on a regular basis, otherwise, it will be necessary to make the appropriate changes in methodology based on and complying with the ISO 29119. In this process, the "organizational test policy" and the "organizational test strategy" will be analyzed. If there are changes, the methodology will be updated through the web platform making it available for all the users.
- **FR-02. Software product development**. In this scenario, the unique actor involved is the analyst-programmer. First, the functional requirements with the customer ought to be elaborated on and analyzed. Thereafter, the functional requirements through the use of the Enterprise Architect (EA) tool ought to be either defined or updated. Next, the functional requirements will be checked through the EA and NDT-Quality tools. If the format is not correct it will be necessary to redefine or update them, otherwise, those will be subjected to customer validation.
- **FR-03. Assess compliance with ISO 29119**. In this scenario, the unique actor involved is the test manager. First, a self-assessment of compliance with the ISO29119 through the web platform must be performed. Then, the evolution of compliance with the standard will be evaluated through the web platform. If the level of compliance is not satisfactory, it will be necessary to make a formal request for changes on the methodology of the organization.
- **FR-04. Software quality management**. In this scenario, the unique actor involved is the test manager. The first step consists of making an analysis of the code's quality through the web platform. Next, the results obtained will be analyzed. If the level of compliance is not satisfactory, a formal request for changes on the project's quality must be done.
- **FR-05. Test execution report**. In this scenario, the unique actor involved is the test manager. First, the available assets must be registered. Then, the environment must be restored to its initial state. After these steps, it is necessary to recollect the learned lessons after the study of the results of the test execution and finally, to report the completion of the test execution.

It is important to note that, to perform this step, the "FR-10. Test execution" must be performed by the testers.

- **FR-06. Test plan monitoring**. In this scenario, two actors are involved: the test manager and the technical manager. The first thing that the test manager has to do is to manage different tasks for configuring the environment and the test plan and later, to set up the environment for monitoring the test plan. During the process of monitoring, several decisions may be taken. The major duties on the part of the test manager are: (i) report on the progress of the test plan, (ii) control the test plan and report control actions to the technical director. On the side of the technical manager has to: (i) access the ALAMEDA web platform for gathering information about control actions and progress of the test plan.
- **FR-07. Test plan creation and maintenance**. In this scenario, the actor "tester" in the role of "tester designer" or an "advanced user" is involved. The first action in this scenario is to understand the context of the project. Once understood, the functional requirements of the project must be checked with through EA and NDT-Quality tools. Then, it is necessary to generate the requirements and acceptance tests documented in the TestLink using the ALAMEDA web platform, which is based on the functional requirements defined previously. The next activities that the user must take are to identify and analyze risks and the treatment that will be applied, determine how long the test plan and its programming will be and finally, record it.
- **FR-08. Design and implement tests**. In this scenario, the actor "tester" in the role of a "tester designer" or an "advanced user" is involved. The first action of this scenario is to generate the test cases using the ALAMEDA web platform. The following actions to be performed: identify feature sets, derive test conditions, derive test coverage items, derive test cases, assemble test sets and finally, derive test procedures. It is important to note that, to perform this step, the "FR-07. Test plan creation and maintenance " must be performed.
- **FR-09. Test environment set-up and maintenance process**. In this scenario, the actor "tester" is involved. For performing this scenario, the test plan and the specification test procedure documents generated in previous steps are necessary. The first action of this scenario is to plan and design the test environment. Then, it has to be properly configured and set up. Finally, it is adamant to prepare data that will be used for the tests. It is important to note that, to perform this step, the "FR-08. Design and implement tests" must be performed.
- **FR-10. Test execution**. In this scenario, the actor "tester" is involved. Through the web platform of ALAMEDA ecosystem, the tester will: (i) carry-out the acceptance tests, (ii) check the code's quality of the project, and (iii) perform the functional and stress tests. From the web platform, the user will register the results for comparing them with older versions. It is important to note that, to perform this step, the "FR-09. Test environment set-up and maintenance process" must be executed.
- **FR-11. Test incident reporting**. In this scenario, the actor "tester" is involved. The first action of this scenario is to analyze the results of the test execution from the web platform. If there are issues, the incident report must be created or updated (where appropriate). It is

http://dx.doi.org/10.5772/66043

important to note that, to carry out this scenario, the "FR-10. Test execution" must be accomplished.

- **FR-12. Users and services management**. In this scenario, the actor "services manager" is involved. The tasks that this actor can perform in this scenario are: to register, to modify or to update users so as to interact with the ALAMEDA ecosystem. These operations can be also applied to the services provided by ALAMEDA.
- **FR-13. Tools set-up management**. In this scenario, the actor "services manager" is involved. The task that this actor can perform in this scenario is to control the different tools that are integrated into the ecosystem for keeping them active and ensuring they work correctly.

3.2. Software tools ecosystem for testing

The software tools that compose the ALAMEDA ecosystem are:

- Functional tests (JUnit, NUnit, Selenium, etc.). These tools depend on a concrete project and its programming language. They are not available from the web platform but it is possible to run their execution (for example, JUnit) through the integrated tools Jenkins or Testlink.
- Load and performance Tests (JMeter). This tool allows the execution of load and performance tests. The execution of these tests can be automated by the use of Jenkins.
- Test management (TestLink). This tool allows the management of tests. The builds of Testlink can be automatically executed from Jenkins.
- Quality software product management (SonarQuBe). The web platform integrates SonarQuBe for analyzing the quality of the code of the product's development.
- Continuous integration and task automation (Jenkins). This tool allows to integrate different tools from other companies not covered in ALAMEDA ecosystem and to automate tasks.

Table 1 shows how the activities of the methodological framework described before and these tools are implemented in ALAMEDA.

Activities	Technological support	Description of technological support
FR-01. Manage methodological guide	Web platform of ALAMEDA ecosystem	Access to the web platform where it will be possible to consult and update the methodological guide
FR-02. Software product development	Desktop platform of ALAMEDA ecosystem	Access to the Enterprise Architect tool where the functional requirements can be defined and generated automated code by NDT-Suite
FR-03. Assess compliance with ISO 29119	Web platform of ALAMEDA ecosystem	Access to the ISO 29119 module where user will: - access to the terms glossary - assess the compliance of the standard - check the status of compliance once evaluated - check the evolution of the results if you have made more than one assessment

Activities	Technological support	Description of technological support
FR-04. Software quality management	Web platform of ALAMEDA ecosystem	Access to the web platform control panel module where user will check the status of SonarQuBe executions
FR-05. Test execution report	Web platform of ALAMEDA ecosystem	Access to the web platform to inform about the test execution report
FR-06. Test plan monitoring	Web platform of ALAMEDA ecosystem - Tools to automate dynamic testing and quality Software	Access to the web platform where the indicators related to the state of the executions or build/s can be displayed
FR-07. Test plan creation and maintenance	- Web platform of ALAMEDA ecosystem - Tools to automate dynamic testing and quality software	Access to the Enterprise Architect and to the web platform where user will generate tests depending on the defined requirements. Access to Testlink for editing and visualizing test plans.
FR-08. Design and Implement Tests	- Web platform of ALAMEDA ecosystem - Tools to automate dynamic testing and quality software	Access to the Enterprise Architect and to the web platform where user will generate tests depending on the defined requirements. Access to Testlink for editing and visualizing test plans.
FR-09. Test Environment Set-Up & Maintenance Process	- Web platform of ALAMEDA ecosystem - Tools to automate dynamic testing and quality software	Access to the web platform where user will: - manage the configuration of ALAMEDA tools - manage the configuration of the continuous integration of Jenkins Access to Jenkins where user will: - configure the plugins installed in Jenkins - configure the automated tasks with Jenkins
FR-10. Test execution	- Web platform of ALAMEDA ecosystem - Tools to automate dynamic testing and quality software	Access to the web platform where user will: - execute load and performance tests Access to Jenkins where user will: - automate the execution process of load and performance (JMeter) and functional tests (JUnit, NUnit, Selenium, etc.). Access to Testlink where user will: - execute tests of the test plan and generate a build (JUnit y JMeter) Access to JMeter where user will: - execute load and performance tests.
FR-11. Test incident reporting	- Web platform of ALAMEDA ecosystem: - Tools to automate dynamic testing and quality software	Access to the web platform where the indicators related to the state of the executions or build/s can be displayed.
FR-12. Users and services management	- Web platform of ALAMEDA ecosystem	Access to the web platform where user sill: - create users - modify users - delete users - manage the different tools provided by the organization
FR-13. Tools set-up management	- Web platform of ALAMEDA ecosystem	Access to the web platform where user sill: - manage the tools configuration. - manage the ISO versions

Table 1. Description of the implementation of the methodological framework.

http://dx.doi.org/10.5772/66043

3.3. ALAMEDA from a user's perspective

In this section, functionalities that can be made through the desktop and the web platform of the ALAMEDA ecosystem are described.

3.3.1. Desktop platform

The desktop platform of ALAMEDA ecosystem consists of an assistant for the creation of functional requirements as a preliminary step to the process of generating automated acceptance tests associated with them.

It is presented as a software layer above deployed EA assisting in the construction of very specific components: actors and system functional requirements. These components are included in EA by previous installation of NDT-Suite and correspond to the basic components to build and validate when creating a specification's requirements model.

The specification's requirements model must follow the structure shown in **Figure 4**.

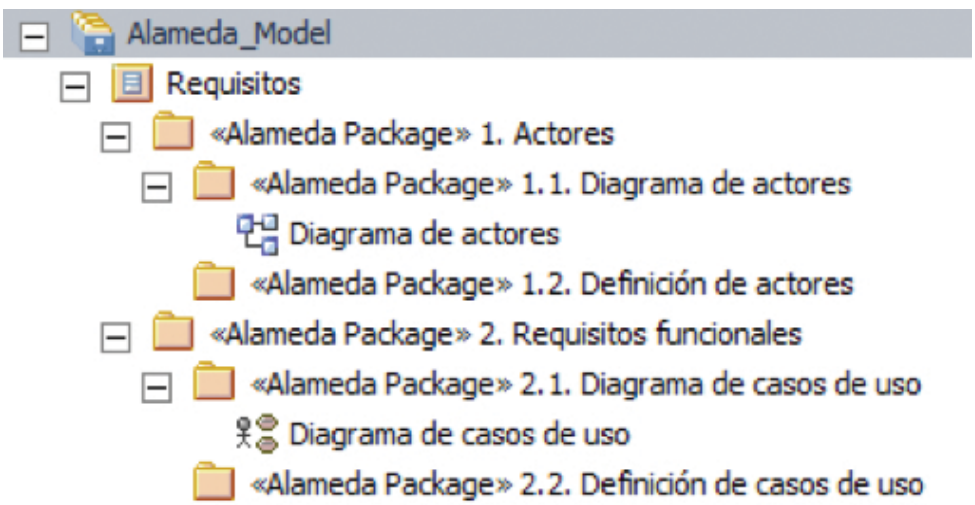

Figure 4. Specification's requirements model.

The construction of the models of functional requirements is based on two operations:

- Collect the functional requirements of a project and the relationship between them and the actors of the system.
- Automatic generation of acceptance tests (provided by ALAMEDA).

Following the previous structure, steps that are necessary to perform are:

- System actor creation/modification. The elements that represent the actors of the system are stereotyped as "AC" and can be found in the EA toolbox provided by NDT-Suite called <NDT System Requirements> DRS Actors. There are two ways of creating these elements:

- Drag the item from the toolbox to the corresponding diagram.
- Select the directory and include the item making right click, choosing the "Add Element" option and selecting the actor item.

- Functional requirements creation/modification. The elements that represent the functional requirements of the system are stereotyped as "FR" and can be found in EA toolbox provided by NDT-Suite called <NDT System Requirements> DRS Functional Requirements. FRs are the most complex elements of the model; they do not just present a definition of a particular value for a specific property, but rather have collections of properties (pre- and post-conditions) and may be specified based on scenarios (sequence of steps) or activity diagrams. The way the RFs are created are in the same way with the actors one. There are different types of relations between the RF and depend on the side of the connection.
  - Extends or include: only applied between RF elements.
  - Use case: only applied between AC and RF elements.

3.3.2. Web platform

The web platform of the ALAMEDA ecosystem brings together the bulk of the functionality of ALAMEDA. In this section, the applicability of each module of this platform is addressed.

- Platform management: for access to the management module it is necessary to access as a manager. The administration panel of the platform (**Figure 5**) displays the following options:
  - New user: it allows registering new users.
  - Edit user: it allows to edit both basic information about users and performs configuration management tools that are assigned to users. Removing users is also allowed.
  - New self-assessment ISO/IEC-29119: it allows the installation of a new version of the ISO/IEC-29119 standard. A new version carries a glossary of terms associated with that role and an associated self-assessment questionnaire.
  - ISO/IEC 29119 management: it allows the management of the versions of the ISO/IEC-29119 standard loaded in the system, which includes removing and updating previous versions.
- Compliance with ISO/IEC-29119: the module in accordance with ISO/IEC-29119 standard (**Figure 6**) has four sections; it depends on an existing version of the standard installed on the platform. Therefore, if there is no version installed on the platform, it will not be possible to access to the other sections of the module. These sections are:
  - Glossary of terms. The glossary of terms shows the collection of terms that defines the norm in the first part for standardizing the terminology used in the field of Testing. In the glossary of terms, each term is shown along with the associated definition.
  - Self-assessment. The self-assessment consists of a series of questions about compliance with the rule that the user must answer based on what is known in reference to his/her

http://dx.doi.org/10.5772/66043

organization. The format shown in the questions are: "Yes", "No" or "Partly", so that each answer carries a score. The questionnaire is broken down into the parts of the standard and within each part, every field and block is referred to the corresponding question. There is an option for storing the state of the questionnaire and complete it later.

- Status of compliance. This section shows the current level of compliance with the standard, corresponding to the last result obtained by the user in a self-assessment. The statistics of the results are broken down into several levels: parts, areas, and blocks.
- Evolution of compliance. This section shows a comparison of the results obtained in the different self-assessments previously completed. The breakdown shown for this option is at the level of parts of the standard.

- Test generation. The test generation module (**Figure 7**) provides all the functionalities associated with the model's requirement specification validation and also, the automation of the creation of acceptance tests. This module is one of the most critical levels of the ALAMEDA ecosystem and it involves the use of three tools: EA, Redmine and TestLink. Two blocks can be distinguished, the one responsible for the tests generation and the one responsible for the tasks generation. The first one allows the creation and management test cases through the use of EA and TestLink and the other one, allows the generation of tasks depending on the unsuccessful test cases managed in the test plan in TestLink. The functionality associated with this module can be divided into four tasks:
  - Tests generation. The test generation process receives as input a specification requirements model in EA format and it returns a test plan generated in the TestLink tool with a build that allows its execution in this tool. The build process begins with a validation phase input model. If errors are found in the model definition, the test generation is not addressed. Instead, a list with the errors found during the process is provided.
  - Generated tests display. The result of the test generation process, regardless of whether they have been generated with errors or not during the validation phase, are displayed in the result viewer.
  - Tasks generation. The process consists of a scanning of the test cases executed in TestLink to generate tasks associated to the Redmine management tool.
  - Tasks monitoring and visualization. The result viewer is similar to the one shown in the test generation module with the difference that the tool deployed is Redmine. The available project tree corresponds to the generated tests. From the generated tests, the viewer can carry out the task monitoring.
- Model generation. The model generation module provides all the functionalities associated with the verification of models and its automated generation for the different phases of the development. This module only uses the EA tool.
- Software and testing analysis. The software analysis module (**Figure 8**) provides functionalities for software analysis and test execution, as well as monitoring the degree of project quality. In this module, the user will:

- measure the quality of the code source of software products through static tests (SonarQube tool).
- execute acceptance tests previously generated in the test execution module (Testlink).
- perform dynamic tests in order to monitor the web nature software product measurement (JMeter).

- Furthermore, as in the test generation module, a result viewer to monitor and track all the conducted activities is provided.
  - Static code analysis. The tool used to perform static code analysis using the web platform is SonaQube. The access form to the execution of static code analysis requests two inputs to perform the test: the project to analyze and its version. The use of compressed projects files as rar or zip format is allowed.
- Control panel. This module allows the monitoring of the elements generated in all the different modules of the web platform.

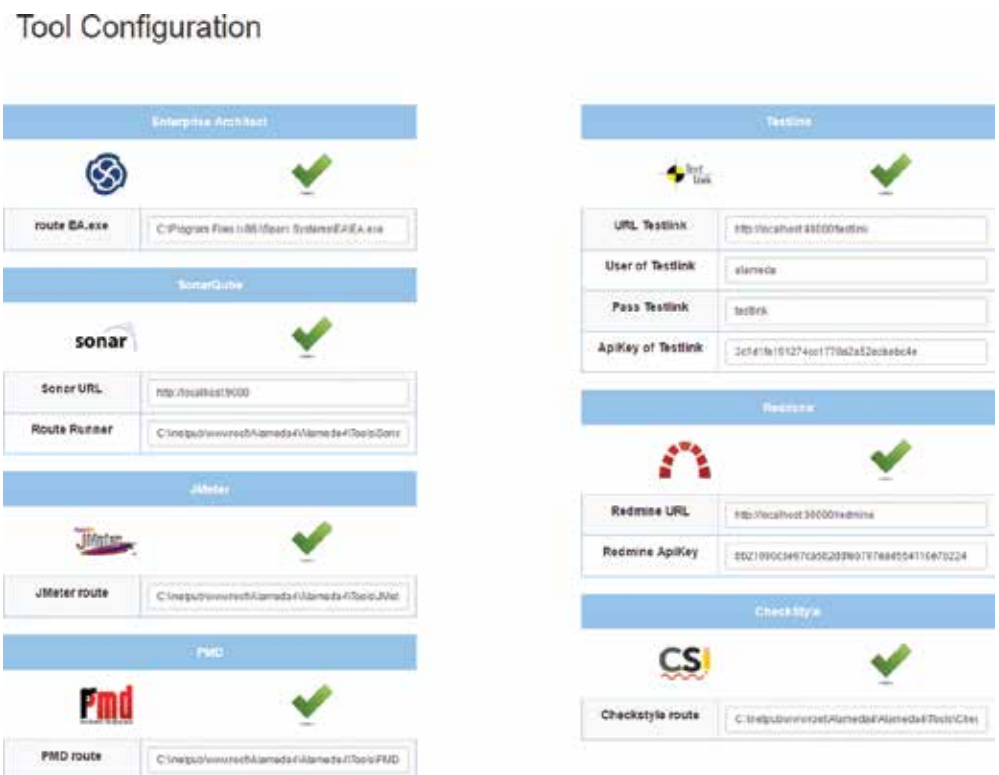

Figure 5. Management platform.

http://dx.doi.org/10.5772/66043

Figure 6. Compliance with ISO/IEC-29119.

Figure 7. Test generation.

Figure 8. Software and testing analysis.

4. Conclusions and future work

This chapter presents the ALAMEDA ecosystem for software testing. Over the last decade, a number of studies, as well as technical and theoretical proposals have emerged to improve the process of software testing. However, the software industry has often carried out this process in a manner that was not standardized and hence, proved difficult to imply to systematization, management and continuous improvement of this process.

The ALAMEDA platform has been developed so as to cover the observed gap in the industry. In this sense, ALAMEDA provides a methodological framework based on the ISO/IEC 29119 standard and the application of the MDE paradigm of software testing (known as MDT).

Regarding the ISO/IEC 29119, this one is a new standard which implies a certain novelty of the software industry. ALAMEDA integrates the main foundations of this standard within its ecosystem in order to facilitate the adoption of this standard by industry. In addition, ALAMEDA provides mechanisms to measure the degree of compliance to the ISO/IEC 29119.Thus, any software company can measure its evolution to meet this standard and carry out a process of continuous improvement in its software development processes.

Regarding the implementation of MDT, ALAMEDA has included the concept of early testing, thanks to it for integrating the NDT methodology within its ecosystem. Thus, it is possible to obtain test cases from the functional requirements and manage these tests effectively. Moreover, this mechanism provides traceability between tests and functional requirements allowing always to know which functionality is being tested.

Finally, and taking into account the features of ALAMEDA, it has not been possible to find a technological solution to cater for all needs associated with the process of software testing from the early stages and complying with the international standard ISO/IEC 29119. In this sense, ALAMEDA improves the state-of-the-art within the management of software quality assurance because it provides an innovative technological solution which is based on well-validated theoretical foundations.

http://dx.doi.org/10.5772/66043

Acknowledgements

This research has been supported by the MeGUS project (TIN2013-46928-C3-3-R), by the SoftPLM Network (TIN2015-71938-REDT) of the Spanish Ministry of Economy and Competitiveness and Fujitsu Laboratories of Europe (FLE). In addition, we would like to acknowledge the leadership of Servinform during the development of the ALAMEDA ecosystem. Our research group has collaborated in the conceptual design of ALAMEDA as well as in analysis, design, and testing tasks of the platform.

Author details

José González Enríquez*, Julián Alberto García-García, Francisco José Domínguez-Mayo and María José Escalona Cuaresma

*Address all correspondence to: jose.gonzalez@iwt2.org

Computer Languages and Systems Department, University of Seville, Seville, Spain

References

[1] Boehm B, Abts C, Brown AW, Chulani S, Clark BK, Horowitz K, Madachy R, Reifer D & Steece B. (2000). Software Cost Estimation with COCOMO II. Prentice Hall PTR, Upper Saddle River, NJ. ISBN 0-13-026692-2.

[2] Jiang J & Klein G. (2000). Software Development Risks to Project Effectiveness. Journal of Systems and Software. 52(1):3–10.

[3] Heckel R & Lohmann M. (2003). Towards Model-Driven Testing. Electronic Notes in Theoretical Computer Science. 82(6):33–43.

[4] Ahmed A. (2012). Software Project Management. A Process-Driven Approach. CRC Press.

[5] Taguchi G, Chowdhury S & Wu Y. Introduction to Design of Experiments. Taguchi's Quality Engineering Handbook, 501–505.

[6] Binder B. (1999). Testing Object-Oriented Systems. Addison Wesley.

[7] RiazAhamed SS. (2010). Studying the Feasibility and Importance of Software Testing: An Analysis. IJEST. 1(3):119–128.

[8] Schmidt DC. (2006). Model-Driven Engineering. Computer-IEEE Computer Society. 39(2):25–31.

[9] Miller J & Mukerji J. (2003). Model Driven Architecture 1.0.1 Guide. Object Management Group, Inc.

[10] Escalona MJ & Aragón G. (2008). NDT. A Model-Driven Approach for Web Requirements. IEEE Transactions on Software Engineering. 34:377–394.

[11] García-García JA, Ortega MA, García-Borgoñon L & Escalona MJ. (2012). NDT-Suite: A Model-Based Suite for the Application of NDT. In Web Engineering. Springer, Berlin Heidelberg. pp. 469–472.

[12] García-García JA, Escalona MJ, Martinez-Garcia A, Parra C & Wojdyeski T. (2015). Clinical Process Management: A Model-Driven & Tool-Based Proposal. In Information Systems Development: Transforming Healthcare through Information Systems.

[13] Tassey G. (2002). The Economic Impacts of Inadequate Infrastructure for Software Testing. National Institute of Standards and Technology, RTI Project, 7007(011).

[14] López AS, Valle DC, Escalona MJ, Lee V & Goto M. (2015). Patient Lifecycle Management: An Approach for Clinical Processes. In Bioinformatics and Biomedical Engineering. Springer International Publishing. pp. 694–700.

[15] Kent S. (2002, May). Model Driven Engineering. In International Conference on Integrated Formal Methods. Springer, Berlin Heidelberg. pp. 286–298.

[16] ISO: ISO/IEC/IEEE 29119 Software Testing. http://www.softwaretestingstandard.org/ (2013). Accessed 10 May 2016.

[17] Fondemenet F & Silaghi R. (2004). Defining Model Driven Engineering Process. 3rd Workshop in Software Model Engineering (WISME2004). October 11–15, Lisbon, Portugal. 2004.

[18] OMG. (2005). UML2.0, Unified Modelling Language. Object Management Group, formal/2005-07-05, 2005.

[19] OMG. (2008). Documents Associated with Meta Object Facility (MOF) 2.0 Query/View/Transformation. Object Management Group. http://www.omg.org/spec/QVT/1.0/. 2008.

[20] Thiry L & Thirion B. (2009). Functional Metamodels for Systems and Software. Journal of Systems and Software. 82(7):1125–1136. DOI: http://dx.doi.org/10.1016/j.jss.2009.01.042. 2009.

[21] Whittle J, Hutchinson J & Rouncefield M. (2014). The State of Practice in Model-Driven Engineering. IEEE Software. 31(3):79–85.

[22] OMG. (2003). MDA Guide of OMG. Version 1.0.1, http://www.omg.org/docs/omg/03-06-01.pdf

[23] Object Management Group (OMG). http://utp.omg.org/. Last accessed: July, 2016

[24] Dai ZR. (2004). Model-Driven Testing with UML 2.0. In Proceedings of the 2nd European Workshop on Model Driven Architecture, Computing Laboratory, University of Kent, Canterbury, UK, 2004, pp. 179–187

[25] Andersin J. (2004). TPI–A Model for Test Process Improvement. In Seminar. University of Helsinki, Helsinki-Finland.

[26] Gutiérrez J, Aragón G, Mejías M, Domínguez F & Cutilla CR (n.d.). Automatic Test Case Generation from Functional Requirements in NDT. WebRE 2012.

[27] García-García JA, Alba M, García-Borgoñon L & Escalona MJ. (2012). NDT-Suite: A Model-Based Suite for the Application of NDT, LNCS 7387, pp. 469–472

[28] García-García JA, Victorio J, García-Borgoñón L, Barcelona MA, Dominguez-Mayo FJ & Escalona MJ. (2013). A Formal Demonstration of NDT-Quality: A Tool for Measuring the Quality using NDT Methodology. The 21st Annual Software Quality Management (SQM) Conference. ISBN 978-0-9563140-8-6.

[29] García-García JA, Cutilla CR, Escalona MJ, Alba M & Torres J. (2011). NDT-Driver, A Java Tool to Support QVT Transformations for NDT. The Twentieth International Conference on Information Systems Development (ISD 2011). ISBN: 978-1-4614-4950-8.

[30] Escalona MJ, Gutiérrez JJ, Mejías M, Aragón G, Ramos I, Torres J & Domínguez FJ. (2011). An Overview on Test Generation from Functional Requirements. Journal of Systems and Software. 84(8):1379–1393, ISSN 0164-1212, http://dx.doi.org/10.1016/j.jss.2011.03.051.

[31] Ostrand TJ & Balcer MJ. (1988). Category-Partition Method. Communications of the ACM. 31(6):676–686.

Chapter 8

Improving Quality Assurance in Multidisciplinary Engineering Environments with Semantic Technologies

Dietmar Winkler, Marta Sabou and Stefan Biffl

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66222

Abstract

In multidisciplinary engineering (MDE) projects, for example, automation systems or manufacturing systems, stakeholders from various disciplines, for example, electrics, mechanics and software, have to collaborate. In industry practice, engineers apply individual and highly specialized tools with strong limitation regarding defect detection in early engineering phases. Experts typically execute reviews with limited tool support which make engineering projects defective and risky. Semantic Web Technologies (SWTs) can help to bridge the gap between heterogeneous sources as foundation for efficient and effective defect detection. Main questions focus on (a) how to bridge gaps between loosely coupled tools and incompatible data models and (b) how SWTs can help to support efficient and effective defect detection in context of engineering process improvement. This chapter describes success-critical requirements for defect detection in MDE and shows how SWTs can provide the foundation for early and efficient defect detection with an adapted review approach. The proposed defect detection framework (DDF) suggests different levels of SWT contributions as a roadmap for engineering process improvement. Two selected industry-related real-life cases show different levels of SWT involvement. Although SWTs have been successfully applied in real-life use cases, SWT applications can be risky if applied without good understanding of success factors and limitations.

Keywords: Semantic Web Technologies, automation systems engineering, multidisciplinary engineering, quality assurance, defect detection, engineering process improvement

1. Introduction

In multidisciplinary engineering (MDE) environments, engineers coming from different disciplines have to collaborate and exchange data. For instance, in automation systems engineering (ASE) projects, mechanical, electrical, and software engineering disciplines work together to build manufacturing plants, power plants, or steel mills [1]. However, engineers typically use isolated tools with data models that provide only limited capabilities regarding data exchange and synchronization of engineering plans across disciplines and organizations [2]. Such loosely coupled tools and incompatible data models hinder efficient engineering processes, project management, quality assurance, and engineering process improvement [3]. Due to limited collaboration capabilities, engineering projects become risky and error-prone.

However, the collaboration of different engineering disciplines requires efficient mechanisms for quality assurance and defect detection [4]. In the context of this paper, defects represent errors in engineering plans and design documents, deviations and inconsistencies between engineering plans. Established tool suites, that is, all-in-one solutions, such as EPlan Engineering Configuration (EEC)[1] or COMOS[2], with integrated data models typically include basic quality assurance mechanisms, such as syntax checks or basic consistency checks to identify intra-disciplinary defects (i.e., defects within one engineering artifact or between artifacts within one discipline). It is worth noting that there is no strong support for identifying interdisciplinary defects, that is, defects that affect two or more disciplines. Thus, in tool networks, where no integrated data models are available, further research is required to enable effective and efficient defect detection. Performed research revealed that there is very limited support for defect detection in heterogeneous environments. The latter include but are not limited to identifying inconsistencies between variables of mechanical and software (control) engineering artifacts because of the technical heterogeneity of tools and the semantic heterogeneity of data models [5].

The application of reviews [6] (or software inspection) in software engineering supports early defect detection [6], typically executed by human experts in a paper-based way, that is, without or with limited tool support. Thus, defect detection requires a high effort and includes risks to oversee important defects, even in homogenous engineering artifacts such as design documents or software code. In distributed and heterogeneous engineering environments, reviews require additional knowledge of human experts, who are sufficiently familiar with at least two related disciplines and require high cognitive skills. Again, review tasks are often executed manually by these experts and require as such considerable effort. It is worth mentioning that preliminary research indicates that they often miss important defects [7], especially if more than one engineering discipline is involved. Thus, an important issue is the heterogeneity of data models, embodied within applied tools.

Semantic Web Technologies (SWTs) can provide concepts for closing the gap between data models for data management, that is, the identification of data entities and knowledge, data

[1] EPlan: www.eplan.de/.

[2] COMOS: w3.siemens.com/mcms/plant-engineering-software/de/Seiten/Default.aspx.

http://dx.doi.org/10.5772/66222

organization, storage, and querying for project analysis purposes [8]. Thus, the application of SWTs can provide the foundation for effective and efficient defect detection mechanisms. However, even without using SWTs, systematic defect detection approaches, such as (software) reviews or inspection [9], can support defect detection processes, by applying guidelines and reading techniques for defect detection support. Thus, the main research challenges focus on (a) providing expert support for identifying defects in an interdisciplinary context more effectively and efficiently and (b) improving defect detection in organizations with or without SWTs. A defect detection framework (DDF) aims at providing an approach to provide the degree of SWT contributions from solution approaches without SWT (lowest level) to a fully supported SWT solution (highest level). For every level, this chapter describes a prototype solution of a real-life use case, developed together with industry and research experts for evaluation purposes.

Thus, the main goals of this chapter focus on (a) providing a defect detection framework (DDF) to support the introduction of SWTs in an organization on various levels of granularity and (b) to provide lessons learned from industry prototypes in context of SWT applications for defect detection in industrial and research real-life use cases. Further, the chapter demonstrates different levels of SWT capabilities in industry use cases and report on lessons learned on the strengths and limitations of these approaches.

The remainder of this paper is structured as follows: Section 2 presents related work on multidisciplinary engineering (MDE), defect detection, and Semantic Web Technologies. Section 3 motivates research issues, and Section 4 presents the solution approach, that is, requirements and capabilities for defect detection in context of MDE and SWT and the defect detection framework with and without SWTs. Section 5 discusses industry (real-life) use cases and prototype implementations regarding defect detection capabilities, and Section 6 discusses qualitative assessment results and lessons learned on the limitations of individual levels of semantic integration.

2. Related work

This section summarizes related work on multidisciplinary engineering (MDE) environments, defect detection, and Semantic Web Technologies as foundation for defect detection tool prototypes in context of the defect detection framework (DDF) with and without SWTs.

2.1. Multidisciplinary engineering environments and automation systems engineering

In *multidisciplinary engineering (MDE)* environments for *automation systems engineering (ASE)*, various stakeholders coming from different disciplines have to collaborate along the automation systems life cycle [1, 10]. Typical examples of such MDE environments for ASE include hydro power plant engineering, steel mill engineering, or industrial production systems. Common to all application areas is the contribution of heterogeneous engineering disciplines, that is, mechanical, electrical, and software engineering, with specialized engineering tools and related specific and heterogeneous data models. However, these expert

tools typically have limitations regarding seamless data exchange and collaboration support of engineers [1, 2, 11], a challenge for effective and efficient defect detection. Defects are typically considered as deviations between engineering plans of heterogeneous disciplines, for example, a mismatch of variable definitions and their usage, missing components, or wrong components. Yet, these deviations could be real defects or required and intended changes, requested by an engineer of one specific discipline. For example, the exchange of a hardware sensor from analogue to a digital (required by the electrical engineer) might result in different numbers of pins that need to be addressed by the software engineer in the control software. This change needs to be identified and analyzed early in the engineering process. Changes that arise late in the engineering project, during the commissioning phase, typically require significantly higher efforts for rework [12], during the commissioning phase.

Figure 1 presents a typical sequential engineering process approach in the ASE domain with parallel engineering activities along the project course. Furthermore, selected and important engineering artifacts of leading engineering disciplines and isolated and distributed quality assurance activities for individual process steps are included in this sample process. In common industry projects, engineers often follow such a simplified sequential engineering process [1, 4], that is, system design, implementation, test and commissioning, and operation (**Figure 1**). Changes and defects can have a critical impact on products, engineering projects, and processes, as they are likely to propagate from early to later stages. Therefore, a need arises for early support of defect detection mechanisms in MDE for ASE projects.

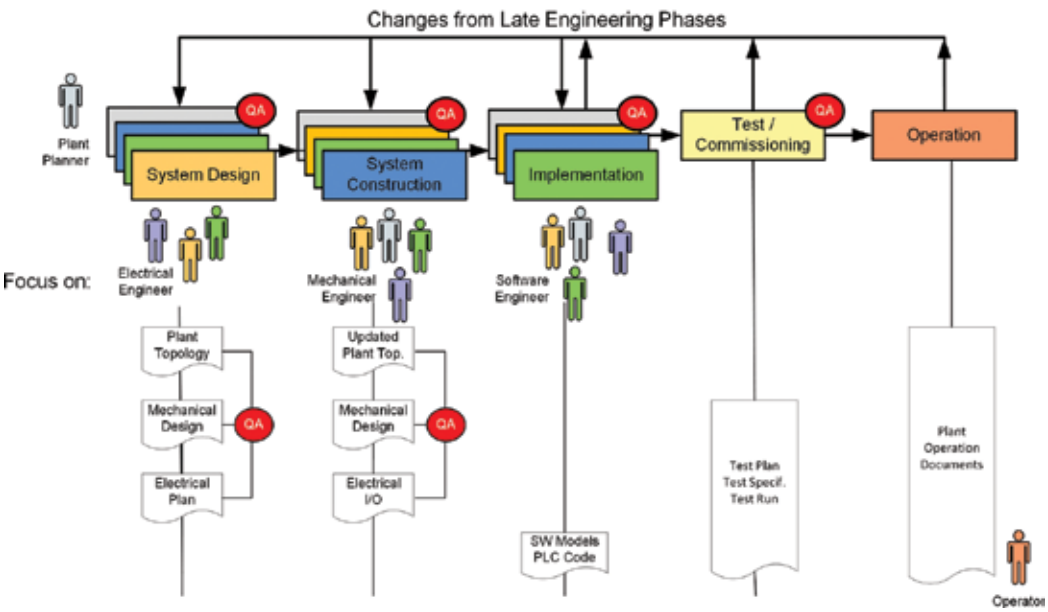

Figure 1. Typical sequential engineering process approach with parallel engineering activities [13].

http://dx.doi.org/10.5772/66222

Synchronization and data exchange in an engineering tool chain are typically executed manually or semi-automatically by applying local supporting tools (e.g., based on local database solutions or spread sheet solutions) [13]. However, these approaches incur high maintenance effort and assume the availability of experts, who are responsible for operating and maintaining these tools (often as add-on activities to their primary work tasks). Thus, there is a need for mechanisms that provide capabilities for mapping heterogeneous data models and for supporting defect detection processes across tools, domains, and data models. To overcome semantically heterogeneous data models, a common data exchange format could bridge the semantic gaps between engineering disciplines. The data format needs to be defined by consideration of all related tools; data to be exchanged need to be mapped to pass information and changes from one discipline (and data model) to the other. To overcome individual effort for data format definition and mapping, a standardized data exchange format is reasonable. *AutomationML* [14–16] is such an emerging data exchange standard that supports efficient synchronization of data produced by different engineering disciplines. However, *AutomationML* represents a data exchange standard and needs additional components that support data exchange, synchronization, and defect detection. The *AML.hub*[3] [17] is a platform for providing efficient data exchange/synchronization based on *AutomationML*. Although the *AML.hub* provides mechanisms for data exchange (e.g., mapping and merging), there is currently no support for quality assurance and defect detection.

2.2. Defect detection in software and systems engineering

Early *defect detection* is a key capability in engineering projects. In software engineering, software reviews are well established to identify defects in software artifacts early and efficiently [9, 18]. These reviews and inspections follow a defect detection process and enable the systematic identification of defects by single reviewers or a review team. Reading techniques are established approaches to support defect detection processes by providing guidelines on how to traverse and read an artifact under review [19]. For example, *scenario-based reading* supports the review of artifacts based on success-critical scenarios or business cases. The main advantage is the focus on real-life use cases that can be prioritized according to engineering risks or business value. *Perspective-based reading* takes the perspectives of different engineering roles, for example, systems architecture, test, or user perspectives to identify defects from various viewpoints. The primary advantage of this approach is the ability of the approach to identify defects from different perspectives based on individual experiences of the experts. In MDE environments, where engineers come from different disciplines, these disciplines may offer useful perspectives for reviewing engineering artifacts to find a variety of different defects.

However, in MDE environments review approaches are often conducted manually by experts. Thus, traditional reviews are effort consuming, expensive, and error-prone because reviewers can easily oversee related defects in engineering plans [20]. Mechanisms with tool-supported defect detection [21] aim to guide reviewers through the review process by focusing on most

[3] AML.hub video: https://www.youtube.com/watch?v=nPg66g46-eM&feature=youtu.be, access: 2016-09.

relevant aspects or critical system characteristics of the engineering artifacts [22]. In software engineering, *Gerrit Code Review*[4] is a widely used review approach for code reviews. Its main features include the comparison of two code variants (e.g., newly available/changed code components and previously developed code component versions, typically stored in a central code repository, GIT[5]), commenting of code fragments, and decision support (e.g., accepting or rejecting modifications). However, *Gerrit Code Review* does not provide any specific guidelines or process support for defect detection in code documents and is focused on software code or applicable on (structured) text elements. In context of defect detection in MDE environments, this approach is comparable to the *Focused Inspection approach* [23] that takes as input deviations or change analysis results between engineering plans to identify changes/defects from the perspective of mechanical, electrical, or software engineers. In order to reduce the manual effort in reviewing activities, SWT-based mechanisms can provide tool support for defect detection. Other tool solutions, such as *DefectRadar*[6], support defect detection by providing annotations and extensive commenting. *DefectRadar* has been applied in building automation without connection to software and/or systems engineering. As such, within the context of review support, the concept of *DefectRadar* can help to better identify entities, relationships, and attributes in different types of documents (e.g., engineering plans or models) to enhance understanding of the artifacts and to support defect detection. To the best of the authors' knowledge, there exists no tool that strongly supports defect detection in MDE environments in a comprehensive way. It is worth noting that Semantic Web Technologies (SWT) represent a promising approach to address these gaps and to support defect detection in MDE environments.

2.3. Semantic Web Technologies for defect detection

Semantic Web Technologies (SWTs) were used in several works to support consistency management across engineering models, including defect detection. Feldmann *et al*. [5, 25] focus on identifying inconsistencies that may arise among diverse engineering plans and engineering models created during the ASE process. Such inconsistency detection contributes to the increased productivity of the engineering process as it supports the detection of potentially severe defects early in the engineering process. The *resource description framework* (RDF) is used to uniformly represent engineering models, which are then queried with SPARQL,[7] a query language for ontologies similar to SQL for databases, to detect defects. Kovalenko *et al*. [26] present an ontology-based approach to automatically detect inconsistencies across heterogeneous engineering data sets. Ontologies are used to explicitly represent the discipline-/tool-specific knowledge and data in a machine-understandable form. Mappings are then defined between the ontologies to model cross-disciplinary (or cross-tool) relations between the data models and data sets explicit for knowledge integration. SPARQL queries are executed over the discipline/tool ontologies regarding the defined mappings in order to perform inconsis-

[4] Google Gerrit: https://www.Gerritcodereview.com, access 2016-09.

[5] GIT: https://git-scm.com/, access 2016-09.

[6] DefectRadar: https://www.defectradar.com/de, accessed 2016-09.

[7] SPARQL: www.w3.org/TR/rdf-sparql-query/, access 2016-09.

http://dx.doi.org/10.5772/66222

tency detection across discipline/tool boundaries. An approach for the automated validation of plant models is presented in [27], where CAEX models are transformed to ontology-based representations and verified through SPARQL queries. In the area of requirements and test case management, Feldmann et al. [28] present a modeling approach that enables the early integration of requirements and test cases. Furthermore, the authors present a case study based on Semantic Web Technologies (SWTs) to ensure the consistency of requirements as well as of requirements and test cases. A conceptual model that describes the main elements for this use case is developed and then formalized as an ontology. Reasoning mechanisms are applied to support various consistency checks-related cases and requirements. Basically, these works show how SWTs can be useful for defect detection in the automation systems domain.

3. Research issues and approach

In the context of *multidisciplinary engineering (MDE)* for *automation systems engineering* (ASE), strong limitations regarding early and efficient *defect detection* for distributed tools and heterogeneous data models were observed, where SWTs can help to semantically integrate these heterogeneous data models. An important question is, therefore, how well SWT mechanisms can help improve defect detection. From this fundamental question, the following research issues are derived:

RI.1: What are success-critical requirements in MDE to enable early, effective, and efficient defect detection to reduce project risks? Based on observations and literature, this chapter will identify a set of success-critical requirements to support managers and engineers in identifying defects early, effective, and efficient.

RI.2: How can SWT contributions be organized in a defect detection framework (DDF) to support defect detection on different levels? Different levels of SWT contributions include defect detection without SWT, defect detection with common concepts (basic approach), and defect detection based on semantically integrated data (advanced approach). However, different levels of SWT contribution can require process changes, additional effort for implementation and application, knowledge experts for SWT implementation, and may need new/adapted methodological approaches and tooling support.

The main question is how this framework can be used to introduce SWT-driven defect detection in organizations and how to establish an appropriate improvement strategy.

RI.3: What are the benefits and limitations of selected industrial cases and prototype implementations with respect to SWT contributions and defect detection in MDE environments? In the context of the CDL-Flex[8] research laboratory, researchers and industry collaborators developed application scenarios and real-life cases that focus on engineering process improvement and defect detection in industry contexts. Selected industry prototypes are presented for every level of the DDF including discussions on the benefits and limitations in context of requirements and

[8] CDL-Flex: Christian Doppler Laboratory for Software Engineering Integration for Flexible Automation Systems, http://cdl.ifs.tuwien.ac.at.

expected capabilities. Thus, RI.3 discusses results of prototype solutions and the strengths and limitations of SWT mechanisms in the context of an MDE development processes.

4. Defect detection framework with/without SWTs

This section summarizes basic requirements, needed defect detection capabilities, and introduces the defect detection framework (DDF) for assessing the level of SWT contributions for defect detection in MDE environments.

4.1. Requirements for defect detection in MDE environments

Requirements and key capabilities have been derived by industry and research experts as a foundation for supporting (a) the development of the defect detection framework and (b) for providing key requirements for offering tool support for defect detection. Basically, requirements include four different core topics: process capabilities (ability to support systematic, traceable, and repeatable processes from quality assurance perspective), organizational requirements (for organizing activities and engineering knowledge), defect detection capabilities (i.e., the core component for defect detection), and tool capabilities that focus on needs for defect detection tool support.

Process capabilities include process support for defect detection, embedded within engineering processes:

- *Systematic defect detection processes* to enable repeatability and traceability of defects and defect detection processes.
- *Traceability* of quality assurance activities focuses on traceability of quality assurance processes, for example, review processes, and results, that is, defects, in various engineering models across disciplines.

Organizational requirements focus on company and organization issues as prerequisites for implementing (and applying) defect detection approaches.

- *Defined roles and responsibilities* for organizing quality assurance activities.
- *Effort for method implementation* is an important factor in context of method implementation because of trade-off considerations of benefits and costs.
- *Knowledge and skills* needed. Within the context of SWT, knowledge experts can be required to support method implementation and maintenance. These expert efforts typically result in additional costs and need to be considered accordingly.

Defect detection performance represents the core capability for defect detection with focus on defect detection performance (i.e., efficiency and effectiveness) and the effort for applying the method/tool.

http://dx.doi.org/10.5772/66222

- *Defect detection effectiveness* refers to the capability of identifying most defects in engineering artifacts. Tool support should also increase the coverage by encompassing all relevant parts of the engineering object.
- *Method application effort* and *defect detection efficiency* refer to the effort for defect detection related to the identified defects (e.g., defects found per time interval). Tool support might also increase defect detection efficiency.

Tool capabilities include basic requirements for a tool solution (on different levels) to support domain experts in the context of defect detection.

- *Tools support* can help to increase defect detection performance, that is, effectiveness and efficiency. However, annotation support can provide additional information, comments on candidate detects for better assessing the correct interpretation of candidate defects.
- *Browsing capabilities* can help to better cross-check information across disciplines and engineering domain borders. Thus, this capability is seen as most valuable to enable identifying relationships within the engineering artifacts.
- *Automation supported difference checks*. Highlighting differences between model versions, that is, related to changes such as added, modified, or removed parts of engineering artifacts, promises to focus on critical changes in and of engineering artifacts.
- *Reporting capabilities* have to provide capabilities to generate reports out of the defect detection process to provide an overview on identified changes/deviations or defects for (quality) management purposes.

The implementation of SWT mechanisms can make data exchange and synchronization of heterogeneous data models more effective and efficient and can enable effective and efficient defect detection.

4.2. Defect detection framework concept

The application of SWT mechanisms requires knowledge engineering capabilities for data management, mapping, and querying. However, the scarce availability of knowledge engineers and the general scarcity of engineering resources [29] in the automation systems engineering (ASE) domain requires also "light-weight" defect detection approaches with and without SWT capabilities. *Light-weight* refers to the application of basic SWT concepts in engineering projects including support for defect detection and without high additional effort for implementing a comprehensive SWT solution for defect detection. **Figure 2** illustrates the *defect detection framework* with different levels of semantic integration approaches.

Defect detection level 0: isolated engineering plans and heterogeneous data sources. Current and traditional MDE approaches take as input heterogeneous artifact sources coming from various disciplines. This defect detection level does not include SWT contributions. Thus, defect detection processes are purely human based and require considerable manual effort for reviewing. Experts have to bridge the semantic gaps between engineering artifacts manually or by applying isolated support solutions that require high effort for application and mainte-

nance [30]. This applies specifically in the case of data model changes. Beyond the effort-consuming manual activities, high cognitive effort is required by experts [31]. The complexity of engineering plans often hides defects very well in various parts of and views on engineering plans. Thus, in MDE, defect detection on this level is often ineffective and always inefficient. While defect detection can be improved by applying reading techniques, such as scenario-based reading or perspective-based reading [32], the high cognitive effort and the manual bridging of semantic gaps still limit defect detection effectiveness and efficiency.

Defect detection level 1: common concepts bridge heterogeneous data models based on commonly used data aspects. Wache *et al*. [33] distinguish between three ontology-based approaches:, *single-ontology, multiple-ontology*, and *hybrid approaches*, to achieve semantic integration, as follows:

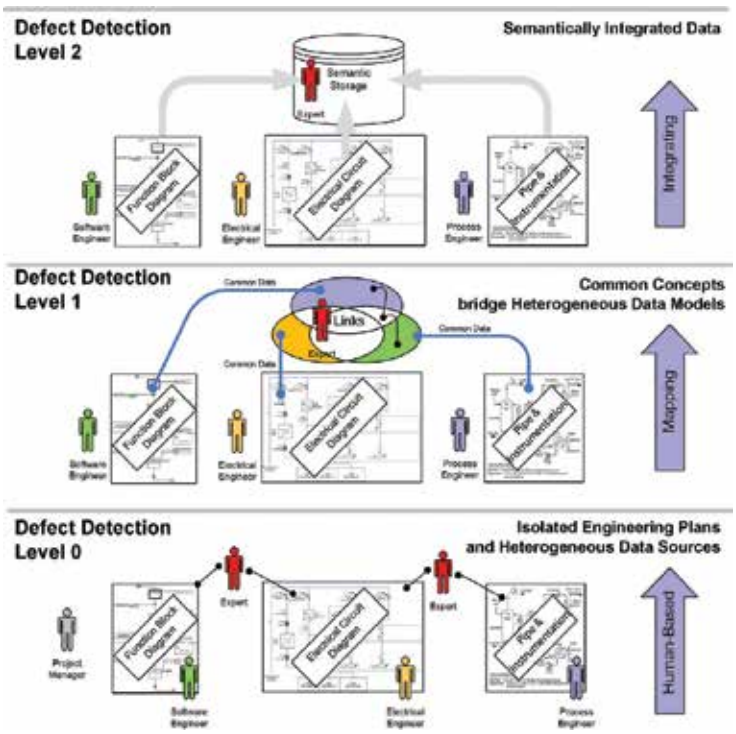

Figure 2. Conceptual approach for the defect detection framework (DDF) in the context of SWT contributions on three levels.

Single-ontology approaches use one ontology as a semantic bridge to integrate several data sources. The challenge is to create an ontology that is sufficiently broad in its coverage to describe all data sources that are integrated. Additionally, an introduction of a new data source requires changes to the ontology.

Multiple-ontology approaches create an ontology to semantically describe each data source to be integrated (i.e., local ontologies) and then create pair-wise mappings between these ontologies.

http://dx.doi.org/10.5772/66222

- *Hybrid approaches* on the other hand combine the previous strategies as follows: Local ontologies are created to describe each data source and are then mapped to a global ontology that contains concepts common to all integrated data sources (i.e., common concepts). The hybrid approach enables engineers to continue working in their well-known environments and enable data synchronization via these common concepts. For instance, in hydro power plant engineering, signals can be used as a common concept. A signal corresponds to electrical wiring or power levels of connectors (electrical discipline), physical layout and position of wires (mechanical discipline), and software variables for control software (software discipline). In the context of defect detection, common concepts enable reviewing engineering aspects by the review team, linked by common concepts. A major advantage of this approach is the lower cognitive burden of reviewers by providing clear links between engineering plans. Tools can be used to automate search tasks during the review by using querying mechanisms. However, this approach requires the upfront investment of defining and maintaining the common concepts data structures.

Defect detection level 2: semantically integrated data. All-in-one solutions (tool suites) typically consist of a common database including plans coming from all engineering disciplines [2] in a homogenous engineering environment. Yet, limitations of that approach include vendor lock-in, limited data exchange capabilities with tools that are not part of the tool suite (common in project consortia [34]), or low flexibility regarding the extensibility of the common data model.

In context of MDE for ASE environments, heterogeneous data models come from different sources and, as such, a solution that assumes a homogeneous data mode is not feasible. In such settings, SWTs can help to provide an integrated view on engineering data from various perspectives. SWT-based engineering model integration aims to bridge semantic gaps in engineering environments between project participants and their tools, who use different local terminologies [35]. Integrated data can then support the analysis, automation, and improvement in MDE processes. Semantic model integration is defined as solving problems originating from the intent to share information across disparate and semantically heterogeneous data [36]. These problems include the matching of data schemes, the detection of duplicate entries, the reconciliation of inconsistencies, and the modeling of complex relations in different data sources [37]. Integrated data and views on engineering data can support defect detection by (a) preparing review packages, that is, scoping specific parts of the overall system for reviewing purposes, and (b) enabling the detection of cross-disciplinary defects, thanks to the explicitly represented links between different disciplines established during the integration process.

While common concepts on level 1 represent manually stated links between data originating from different disciplines, links on level 2 are established semi-automatically by using explicit engineering knowledge. While the semantic integration with SWTs already supports defect detection activities, SWTs can be employed to create defect detection mechanisms, in particular, through semantic querying with SPARQL. Benefits of these approaches are: (a) defect detection tasks are explicitly represented in SPARQL queries, thus allowing periodic query repetition; (b) since queries typically address the ontology concepts, the checks remain valid even if the underlying data model, for example, of signal concepts, changes; and (c) semantic

query languages rely on reasoning mechanisms and are, therefore, well fitted [38] to check model consistency and coherency.

5. Industry use cases and lessons learned

This section presents two key *Use Cases* (UCs) developed in cooperation with industry partners and prototype solutions in context of the defect detection framework (DDF) without/with Semantic Web Technologies.

5.1. Defect detection without Semantic Web Technologies (level 0)

On Level 0, where a set of isolated engineering plans based on heterogeneous data sources is available, no SWTs are used. Engineering knowledge is implicitly embedded within domain experts [24], who are familiar with at least two engineering disciplines, for example, electrical and software or electrical and process discipline, as depicted in **Figure 2**. These experts bridge the semantic gap between aforementioned engineering disciplines manually. Thus, defect detection can become time-consuming and risky because experts can oversee defects easily within a huge amount of given data. For instance, in the hydro power plant domain, 40.000 data points (i.e., signals) are available and need to be reviewed/inspected by experts. Tool solutions like *Gerrit* can be used to compare different engineering model/plan versions to highlight changes to better drive the review process. However, the availability of a textual representation is required for enabling *Gerrit* applications. If no textual documents are available but engineering plans (such as diagrams) are present, *DefectRadar* can be used for annotation and review purposes [39].

Independent of applied tools, guidelines from Review/Inspection, such as perspective-based reading, scenario-based reading, or checklist-based reading technique approaches, can be used to focus on perspectives (defect detection from experts coming from specific disciplines) or scenarios (focus on use cases and application scenarios). Checklists based on the ISO 25010 standard [40] can represent a complementary approach to systematically check important (predefined and application-specific) items to support defect detection. Although such guidelines support review processes, the review itself is still performed manually and often paper based [41].

5.2. Common concepts with limited Semantic Web Technologies (level 1)

The first (basic) level of SWT contributions includes the application of common concepts as a hybrid ontology (see Section 4.2). The main idea of this approach is to elicit commonly used data from involved disciplines (even if they have different terminologies), transform them to a common concept, and map these common data between different disciplines. Thus, specifically used data are left within local tool solutions, and common concepts are centrally stored and are used for mapping data elements between related disciplines. Thus, changes in one discipline can be passed efficiently via common concepts to related disciplines by highlighting changes. Hence, the focus for review is based on deviations and changes (similar to

http://dx.doi.org/10.5772/66222

the *Gerrit* approach). **Figure 3** presents a simple workflow for difference/defect detection approach, and **Figure 4** shows a screenshot of the prototype implementation (i.e., highlighting deviations/changes of different engineering model/plan version as input for review).

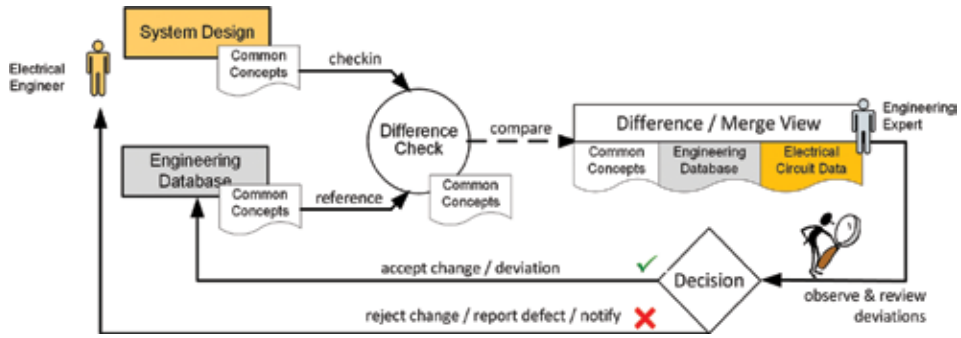

Figure 3. Basic workflow for difference/defect detection with focused reviews.

This concept has been implemented and evaluated as *"Engineering Object Change Management"* (UC 1) in the hydro power plant domain [42]. In the context of this use case *signals* can be seen as the foundation for common concepts that enable efficient data exchange between engineering disciplines and synchronization of heterogeneous engineering plans. A *Virtual Common Data Model* (VCDM) has been elicited with industry experts to link individual disciplines [43, 44]. Note that key parts of signals (coming from different disciplines) have to be identified, transformed to the VCDM, and mapped to enable a seamless data integration. A database holds the VCDM data and provides links to discipline-specific tools and data models. Based on this VCDM concept, a change management process supported by a web-based tool has been designed at the industry partner.

Figure 4. Prototype screenshot of difference checks of the prototype implementation.

The prototype tool is capable of handling changes, that is, new, modified, or removed signals across disciplines and domain borders. A *difference view*, that is, a comparison of already available signals and modified signals, enabled engineers to assess the signal data and to decide whether the deviation should be accepted as a change or be classified as a defect. The main results of the pilot study [4] were that the prototype solution enabled (a) more effective and

efficient data exchange in all phases of the engineering process and (b) observation capabilities of the change management process for project management and control.

In the context of defect detection, the explicit view on "deviation types" (i.e., changes or defects) provides the foundation for a more effective and efficient review process of signal data. An important benefit of this approach is the focus on the most critical deviations from a prior data version instead of reviewing the overall set of engineering data (signals) of the hydro power plant. Note that a typical hydro power plant includes a set of up to 40.000 signals and a small subset (2–5%) of deviations that allow to focus the review to 800 to 2.000 deviating signals. Although the prototype solution provides support for defect detection by presenting differences and deviations to the reviewer, the review process is executed by a human expert team without any specific guidance for defect detection. In this context, guidance refers to applying checklists, scenarios, or perspectives to identify defects more effective and efficient. Justification of defects typically relies on expert domain knowledge and engineering experience and is based on industry and organization best practices.

Although the automated identification of deviations and changes has been found useful by experts in context of the prototype implementation, additional guidelines, such as reading techniques, can help improve defect detection performance in terms of increased effectiveness (i.e., increased number of defects found) and efficiency (i.e., identifying more defects per time interval), and decreased false positives (i.e., wrongly reported defects).

Based on the provided focus on deviations and changes, the *"Focused Inspection"* [45] and "*Model Quality Assurance*" approaches [46] help to improve the unguided defect detection process by supporting (a) the review team selection and (b) their reading process. Individual engineers focus on defined aspects of the systems, for example, on mechanical, electrical, or software characteristics of the engineering plan, and their overlaps with partner disciplines. Thus, the reading technique can take into account individual disciplines, perspectives, and typical candidate defects that occur in discipline-specific engineering plans and across engineering plans. In this use case, *common concepts* (and the VCDM) are used as vehicle for data synchronization and for improving reviewing processes. The reviewing process is supported by a difference view (tool support) and reading technique (method support). In the pilot study [4], both the defect detection effectiveness and efficiency increased. However, an additional knowledge experts or key users that are familiar with knowledge engineering are required for supporting the common concept elicitation, transformation, and mapping process. In the prototype solution, a knowledge expert, that is, one of the authors, supports these process steps. However, additional tool support is planned as future work that guides domain experts and key users to conduct these knowledge management tasks without deep understanding of SWT mechanisms.

5.3. Integrated data with advanced Semantic Web Technologies (level 2)

The second (advanced) level of SWT contributions includes the establishment of integrated data based on a commonly used data exchange language (see Section 4.3). The *AutomationML*

http://dx.doi.org/10.5772/66222

Analyzer[9] (UC 2) uses SWTs to provide an interface for analyzing data from integrated *AutomationML* files.

AutomationML [14] is an emerging data format for exchanging engineering data [47]. While the concerted use of *AutomationML* in an engineering project makes the integration between data from different engineering disciplines easier, tools are still needed for navigating and analyzing integrated *AutomationML* data more easily. To that end, the *AutomationML Analyzer* [48] uses ontology-based technologies to integrate *AutomationML* data to provide easy navigation support within the *AutomationML* data as well as to detect project level inconsistencies and defects through SPARQL querying of the integrated data. The main SWT capabilities used are (a) semantic modeling of an *AutomationML* ontology, which enables semantically enriching the input data; (b) browsing and exploration of the semantic data through *Linked Data*-based mechanisms; and (c) the use of reasoning mechanisms as part of the SPARQL querying activities.

This prototype fits Level 2 of the framework because it uses SWTs both for data integration and for defect detection activities. **Figure 5** presents the concept for the *AutomationML Analyzer*, including review process support. Core components of the concept include the *AutomationML Analyzer* component, providing integrated data and the *review process support* component that enables required review functionality such as browsing through the plant topology, automated execution of defined queries to identify deviations and changes, and querying capabilities for reporting. As a prerequisite, engineering data are available in *AutomationML* data format. However, the concept is applicable to any structured data formats required within a project. In the context of this prototype, the *AML.hub* [17, 49] represents the platform for data synchronization (i.e., mapping and merging of data coming from heterogeneous sources).

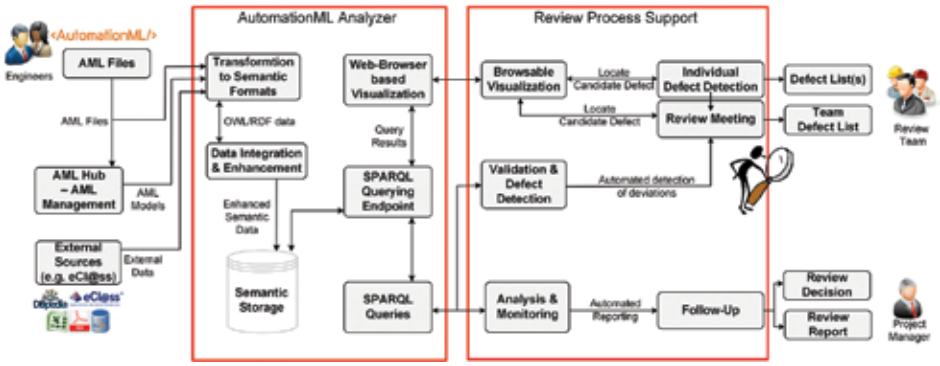

Figure 5. *AML Analyzer* concept for review support [48].

For establishing the *AutomationML Analyzer*(and the underlying mechanisms), a knowledge engineer is required to perform a multitude of tasks, such as transformations, data integration,

[9] AutomationML Analyzer: http://data.ifs.tuwien.ac.at/aml/analyzer, access 2016-09.

and querying capabilities. However, in common industry projects where similar projects are conducted, this effort represents initial activities and the results, for example, transformation, mapping strategies, and queries, can be reused for similar projects within the organization.

6. Discussion, limitations, and future work

The main goal of this paper focuses on building a platform for users and domain experts, that is, experts in MDE and ASE environments, for executing review activities more effectively and efficiently (process improvement) and to improve *defect detection* performance in terms of increased defect detection effectiveness and efficiency (method application improvement). Several strategies can help to improve defect detection processes and activities by using a certain degree of Semantic Web Technology involvement. The defect detection framework (DDF) provides a basic framework for approaches in context of defect detection with a different degree of SWT contributions.

Table 1 presents *identified requirements and expected key capabilities* (y-axis) for tool solutions (RI.1). These requirements and capabilities have been elicited in collaboration with industry and research experts (Section 4.1). RI.2 focuses on the development a *defect detection framework (DDF)*, illustrated in **Figure 2** and represented by the x-axis in **Table 1**, which is capable of providing an overview of different levels of SWT contributions for defect detection. Within the framework of this chapter, three levels of SWT contributions have been found useful for assessing the level of SWT contribution and for driving improvement initiatives in context of quality assurance and defect detection. On different level of SWT contributions, selected industry real-life prototypes have been developed and conceptually evaluated (RI.3) to elicit benefits and limitations of individual SWT-related solutions. Domain experts bridge technical gaps between isolated tools and semantic gaps between heterogeneous data models manually with limited tool support. In addition, defect detection is executed manually with high cognitive effort. To overcome these limitations (i.e., high effort and error-proneness of manual defect detection), common concepts help to bridge the gap based on commonly available data. *Defect Detection Level 1* focuses on these common concepts, that is, commonly used data that support data exchange between disciplines. Transformations and mappings of local tool data to common concepts help to bridge the semantic gaps of individual data. Based on these mappings, changes can be propagated via common concepts from one discipline to the other. Furthermore, defect detection is supported by providing difference/deviation checks to focus on differences between artifact versions or artifacts derived from different disciplines (see **Figure 4** in Section 5.2). This approach has been successfully implemented in industry (see Section 5.2) which supports the engineering team *Defect Detection Level 2*, uses advanced approaches from SWTs, and builds on semantically integrated data that enable browsing, querying, and automation-supported defect detection and reporting—implemented in the *AutomationML Analyzer* (see Section 5.3). Note that this approach uses *AutomationML*, an emerging standard for data exchange in MDE environments for enabling data management, that is, data mapping and synchronization. Yet, the conceptual approach supports various types of data formats and has been implemented with *AutomationML* data (see Section 5.3).

http://dx.doi.org/10.5772/66222

	Level 0 (human based)	Level 1 (common concepts)	Level 2 (integrated data)
Process capability			
Systematic defect detection	Low	High	High
Traceability	Low	High	High
Organizational requirements			
Defined roles and responsibilities	Medium	Medium	Medium
Effort for method implementation	Low	Medium	High
Knowledge and skills needed	Medium	High	High
Defect detection performance			
Defect detection effectiveness	Low	High	High
Defect detection efficiency	Low	High	High
Method application effort	High	Medium	Low
Tool capabilities			
Tool support	Low	Medium	High
Browsing capabilities	Low	Low	High
Automated supported difference checks	Low*	High	High
Reporting capabilities	Low	Medium	High

*Depending on the involved alternative tools, for example, *Gerrit* or *DefectRadar*.

Table 1. Conceptual evaluation of the defect detection framework.

However, to enable the elicitation of common concepts, transformation and mapping (for Level 1) and the construction of integrated data (for Level 2) requires knowledge engineers who support users and domain experts, elicit, and implement SWT approaches.

Table 1 presents an assessment of SWT contribution levels (and prototype solutions) and identified requirements for process and tool support in context of defect detection. *Human-based defect detection* (Level 0) is based on expert activities and their expertise with limited process support and limited tool support. However, due to organizational activities, reviews and inspection can still be implemented in MDE environments on a basic level (e.g., by using organizational guidelines and supporting tools). *Common Concepts* (Level 1) support defect detection by enabling links and mechanisms for data synchronization and defect detection, for example, highlighting changes in artifacts derived from different disciplines or artifact versions, to focus on recent changes and deviations. Instead of reviewing thousands of engineering data, they can focus on deviations, i.e., changes, inconsistencies, or defects. In the evaluation context, the prototypes showed:

- improved process and organizational support,
- improved defect detection capabilities, and

- some basic tool support.

Integrated Data (Level 2) builds on a high degree of SWT and data integration and enables advanced defect detection by also supporting review processes, querying, navigation, and enhanced tool support.

With focus on the *defect detection framework* (RI.2), the three-level concept has considerable benefits for classifying SWT contributions as starting point for improving/introducing SWTs in an organization to improve collaboration and defect detection performance. On the other hand, defect detection performance can also be improved with low effort, by introducing systematic reading techniques or tool support to focus on most critical and important deviations. Nonetheless, to benefit from SWTs capabilities, for example, queries, reporting, and navigation, knowledge engineering capabilities are required to support engineers in data management and query definition.

6.1. Limitations and threats to validity

The *Requirements and Tool Capabilities* have been detailed and discussed with representative industry and research experts in the field of automation systems engineering. These industry and research experts have been recruited from the automation systems development domains, for example, the hydro power plant domain. Although identified requirements and capabilities match to the needs of these experts, additional requirements and needed tool capabilities might arise if different application domains are considered, for example, building automation, factory automation, or chemical production systems. For instance, common concepts need to be adapted according to available data in related engineering domains.

The *defect detection framework (DDF)* has been developed within the context of the current research work. It has been executed in the CDL-Flex, which is driven by industry partners and their respective individual needs. However, it is the view of the authors that the framework may assist industry and research entities so as to drive and foster improvement initiatives. Although this basic framework has been found useful in the evaluation context, additional levels of SWT contributions might be reasonable and applicable for different application domains. Further research work is needed to include additional applied SWT contributions applicable in different domains.

Selected *Use Cases and Prototype Solutions* have been developed based on industry-specific requirements that have been derived from industry partners in the automation systems domain, mainly hydro power plant engineering, steel mill engineering, or production automation engineering. Assuming comparable requirements in different domains, SWT contributions need to be evaluated in these additional application domains. Nonetheless, the conceptual real-life use cases and prototype implementations have been evaluated in various industry contexts successfully (see Section 4 for the solution concepts and Section 5 for individual industry real-life user cases). Based on described prototype solutions, additional application experiences in the defined industry context (i.e., business domain of industry partners) are needed. In addition, requirements from related business domains need to be explored and considered in the presented solution concepts. Therefore, the plans include more

http://dx.doi.org/10.5772/66222

detailed large-scale empirical studies (e.g., case studies) in real-world industry settings at existing industry partners and beyond. SWTs can make defect detection more effective and efficient. It is worth noting though that a higher level of SWT contribution requires considerable knowledge engineering capabilities to support data management and query definition. Although SWT concepts have been successfully applied in industry, their application can be risky without good understanding of their success factors and limitations.

6.2. Future work

Future work focuses on evaluating and improving the presented solution concepts and prototype solutions, extending and enhancing the prototype solution within the current application domains and addressing needs from additional application domains, and extending the platform toward a comprehensive process support for enabling and improving comprehensive reviewing with SWT and tool support. More specifically, future work includes:

- Improving solution concepts and prototype solutions based on experiences from real-world applications, for example, pilot application in real-life industry projects at target industry partners from the hydro power plant domain, steel mill engineering, and manufacturing system development. Future goals also focus on exploring requirements and needs in additional industry application domains for upcoming development steps.
- Prototype evaluation will be extended toward more detailed investigations in large-scale industry contexts with large-scale real-world data (beyond pilot applications and prototype applications in small contexts).
- Future plans also include strengthening the process support for defect detection by providing a more comprehensive view on quality assurance based on established review and inspection processes (see Section 2.2 for the traditional review process approach).
- Finally, the defect detection framework with SWT contributions needs to be revisited, improved, and extended to establish a framework to drive engineering process improvement with SWT mechanisms.

It is the authors' view that SWTs can make defect detection more effective and efficient. It is worth noting though that a higher level of SWT contribution requires considerable knowledge engineering capabilities to support data management and query definition. Although SWT concepts have been successfully applied in industry, their application can be risky without good understanding of their success factors and limitations [50].

Acknowledgements

Part of this work was supported by the Christian Doppler Forschungsgesellschaft, the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development, Austria.

Author details

Dietmar Winkler[1,2*], Marta Sabou[2] and Stefan Biffl[2]

*Address all correspondence to: dwinkler@sba-research.org

1 SBA Research gGmbH & Institute of Software Technology and Interactive Systems, CDL-Flex, Vienna University of Technology, Vienna, Austria

2 Institute of Software Technology and Interactive Systems, CDL-Flex, Vienna University of Technology, Vienna, Austria

References

[1] Biffl S., Schatten A., Zoitl A. Integration of heterogeneous engineering environments for the automation systems lifecycle. In: Proceedings of the 7th International Conference on Industrial Informatics (INDIN); 23–26 June 2009; Cardiff, Wales. IEEE; 2009. p. 576–581. doi: 10.1109/INDIN.2009.5195867

[2] Fay A., Biffl S., Winkler D., Drath R., Barth M. A Method to evaluate the openness of automation tools for increased interoperability. In: Proceedings of the 39th Annual Conference of the IEEE Industrial Electronics Society (IECON); 10–13 November; IEEE; 2013. doi: 10.1109/IECON.2013.6700266

[3] Biffl S., Lüder A., Winkler D. Multi-disciplinary engineering for Industrie 4.0: semantic challenges, needs, and capabilities. In: Biffl S., Sabou M., editors. Semantic Web for Intelligent Engineering Applications. 1st ed. Springer, New York; 2016.

[4] Winkler D., Moser T., Mordinyi R., Sunindyo W.D., Biffl S. Engineering object change management process observation in distributed automation systems projects. In: Industrial Proceedings of the 18th EuroSPI Conference; Roskilde, Denmark; 2013.

[5] Feldmann S., Herzig S.J.I., Kernschmidt K., Wolfenstetter T., Kammerl D., Qamar A., Lindemann U., Krcmar H., Paredis C.J.J., Vogel-Heuser B. Towards effective management of inconsistencies in model-based engineering of automated production systems. In: Proceedings of the IFAC Symposium on Information Control in Manufacturing (INCOM); 2015.

[6] Rössler P. Schlich M., Kneuper R. Reviews in der System- und Softwareentwicklung: Grundlagen, Praxis, kontinuierliche Verbesserung. 1st ed. dpunkt.verlag; 213. 176 p.

[7] Kovalenko O., Serral E., Sabou M., Ekaputra F.J., Winkler D., Biffl S. Automating cross-disciplinary defect detection in multi-disciplinary engineering environments. In: Proceedings of 19th International Conference on Knowledge Engineering and Knowledge Management (EKAW); Linköping, Sweden. 2014.

[8] Biffl S., Sabou M., editors. Semantic Web for Intelligent Engineering Applications. 1st ed. Springer, New York; 2016.

[9] Aurum A., Petersson H., Wohlin C. State-of-the-art: software inspections after 25 years. Software Testing, Verification and Reliability. 2002;12(3):133–154.

[10] Biffl S., Moser T., Winkler D. Risk assessment in multi-disciplinary (Software+) engineering projects. International Journal of Software Engineering and Knowledge Engineering (IJSEKE), Special Session on Risk Assessment. 2011;21(2):211–236.

[11] Barth M., Drath R., Fay A., Zimmer F., Eckert K. Evaluation of the openness of automation tools for interoperability in engineering tool chains. In: Proceedings of the 17th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA; 17.-21.09.2012); Krakow, Poland. IEEE; 2012.

[12] Winkler D., Biffl S. Improving quality assurance in automation systems development projects. In: M. Savsar, editor. Quality Assurance and Management. Intech Publishing; 2012. p. 379–398. doi: 10.5772/33487

[13] Biffl S., Mordinyi R., Steininger H., Winkler D. Integrationsplattform für anlagenmodellorientiertes Engineering - Bedarfe und Lösungsansätze. In: Vogel-Heuser B., Bauernhansl T., tenHompel M., editors. HandbuchIndustrie 4.0. 2nd ed; 2016.

[14] IEC 62714-1: Engineering data exchange format for use in industrial automation systems engineering—Automation Markup Language—Part 1: Architecture and general requirement. International Standard; 2014

[15] IEC 62714-2: Engineering data exchange format for use in industrial automation systems engineering—Automation Markup Language—Part 2: Role class libraries. International Standard; 2015

[16] IEC 62714-3: Engineering data exchange format for use in industrial automation systems engineering—Automation Markup Language—Part 3: Geometry and Kinematics. International Standard; 2017 (forthcoming)

[17] Winkler D., Biffl S., Steininger H. Integration von heterogenen Engineering Daten mit AutomationML und dem AML.hub: Konsistente Daten über Fachbereichsgrenzen hinweg. develop3. 2015;3/15.

[18] Laitenberger O., DeBaud J-M. An encompassing life cycle centric survey of software inspection. Journal of Systems and Software. 2000;50(1):5–31.

[19] Travassos G., Shull F., Fredericks M., Basili V.R. Detecting defects in object-oriented designs: using reading techniques to increase software quality. ACM Sigplan. 1999;34(10):47–56.

[20] Biffl S., Halling M. Investigating the defect detection effectiveness and cost benefit of nominal inspection teams. IEEE Transactions on Software Engineering. 2000;29(5):385–397.

[21] Lanubile F., Mallardo T., Calefato F. Tool support for geographically dispersed inspection teams. Software Process Improvement and Practice. 2003;8(4):217–231.

[22] Anderson P., Reps T., Teitelbaum T., Zarins M. Tool support for fine-grained software inspection. IEEE Software. 2003;20(4):42–50.

[23] Winkler D., Biffl S. Focused inspections to support defect detection in multi-disciplinary engineering environments. In: Proceedings of the 16th International Conference on Product-Focused Software Process Improvement (PROFES); Bozen-Bolzano, Italy. Springer, New York; 2015.

[24] Winkler D., Musil J., Musil A., Biffl S. Collective intelligence-based quality assurance: combining inspection and risk assessment to support process improvement in multi-disciplinary engineering. In: Proceedings of the 23rd EuroSPI Conference; 14.-16.09.2016; Graz, Austria. Springer, New York; 2016. 163–175 p.

[25] Feldmann S., Herzig S.J., Kernschmidt K.,Wolfenstetter T., Kammerl D., Qamar A., Lindemann U., Krcmar H., Paredis C., Vogel-Heuser B. A comparison of inconsistency management approaches using a mechatronic manufacturing system design case study. In: Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE). IEEE; 2015.

[26] Kovalenko O., Moser T. Using explicit and machine-understandable engineering knowledge for defect detection in automation systems engineering. In: Proceedings of International Doctoral Symposium on Software Engineering and Advanced Applications (IDoSEAA); Oulu, Finland; 2011.

[27] Abele L., Legat C., Grimm S., Müller A.W. Ontology-based validation of plant models. In: 11th IEEE International Conference on Industrial Informatics (INDIN); IEEE; 2013. 236–241 p.

[28] Feldmann S., Rösch S., Legat C., Vogel-Heuser B. keeping requirements and test cases consistent: towards an ontology-based approach. In: 12th IEEE International Conference on Industrial Informatics (INDIN); 2014. 726–732 p.

[29] Ekaputra F., Serral E., Winkler D., Biffl S. An analysis framework for ontology querying tools. In: Proceedings of the 9th International Conference on Semantic Systems (I-SEMANTICS) in conjunction with the 13th International Conference on Knowledge Management and Knowledge Technologies (I-KNOW). ACM; 2013.

[30] Biffl S., Mordinyi R., Moser T. Anforderungsanalyse für das integrierte Engineering – Mechanismen und Bedarfe aus der Praxis. ATP Edition; 2012;54(5):28–35.

[31] Blackwell A., Green T. Notational systems—the cognitive dimensions of notations framework. In: Carroll J., editor. HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science.Morgan Kaufmann, San Francisco; 2003. 103–134 p.

http://dx.doi.org/10.5772/66222

[32] Winkler D. Improvement of Defect Detection with Software Inspection Variants: A Large-Scale Empirical Study on Reading Techniques and Experience. VDM. Saarbrücken, Germany, 2008.

[33] Wache H., Vögele T., Visser U., Stuckenschmidt H., Schuster G., Neumann H., Hübner S. Ontology-based integration of information—a survey of existing approaches. In: Proceedings of IJCAI workshop: ontologies and information sharing; 2001. pp. 108–117.

[34] Ekaputra F., Serral E., Winkler D., Biffl S. Addressing data integration and communication issues in project consortia. In: Proceedings of the International Conference on Data and Software Engineering (ICoDSE), Bandung, Indonesia; 2014.

[35] Aldred L., van der Aalst W., Dumas M., Hofstede A. Understanding the challenges in getting together: the semantics of decoupling in middleware. BPM Center Report. 2006;BPM-06-19.

[36] Halevy A. Why your data won't mix. Queue. 2005;3(8):50–58.

[37] Noy N.F., Doan A.H., Halevy A.Y. Semantic integration. AI Magazine. 2005;26(1):7–9.

[38] Moser T., Mordinyi R., Mikula A., Biffl S. Efficient integration of complex information systems in the ATM domain with explicit expert knowledge models. Complex Intelligent Systems and their Applications. 2010;41:1–19.

[39] Winkler D., Biffl S. Collaborative Model Review Support for AutomationML. 4th AutomationML User Conference, Esslingen, Germany, 2016 (upcoming).

[40] ISO/IEC 25010:2011. Systems and Software Engineering—Systems and Software Quality Requirements and Evaluation (SQuaRE)—System and Software Quality Models. Standard, March 2011.

[41] Melo W., Shull F., Travassos G. Software Review Guidelines. Technical Report ES-556/01, COPPE/UFRJ, 2001.

[42] Winkler D., Biffl S. Focused inspection to support defect detection in automation systems engineering environments. Technical Report. TU Wien. IFS-CDL 15-02. Sep 2015. Online available: http://qse.ifs.tuwien.ac.at/publication/IFS-CDL-15-02.pdf.

[43] Sunindyo W., Moser T., Winkler D. Project progress and risk monitoring in automation systems engineering. In: Proceedings of the 5th Software Quality Days. Springer, New York; 2013.

[44] Sabou M., Kovalenko O., Novak P. Semantic modelling and acquisition of engineering knowledge. In: Biffl S., Sabou M., editors. Semantic Web for Intelligent Engineering Applications. 1st ed. Springer, New York; 2016.

[45] Winkler D., Ekaputra F.J., Biffl S. AutomationML review support in multi-disciplinary engineering environments. In: Proceedings of the 21st IEEE International Conference on Emerging Technologies and Factory Automation (ETFA); Berlin, Germany. IEEE; 2016.

[46] Winkler D., Wimmer M., Biffl S. Model quality assurance for multi-disciplinary engineering. In: Biffl S., Lüder A., Gerhard D. (eds): Multi-Disciplinary Engineering of Cyber-Physical Production Systems, Book Chapter, Chapter 17, 2017 (upcoming).

[47] Drath R., editor. Datenaustausch in der Anlagenplanung mit AutomationML: Integration von CAEX, PLCopen XML und COLLADA. 1st ed. Springer; 2010.

[48] Sabou M., Ekaputra F.J., Kovalenko O. Supporting the engineering of cyber-physical production systems with the AutomationML analyzer. In: Proceedings of the CPPS Workshop; Vienna, Austria. 2016.

[49] Mordinyi R., Wimmer M., Biffl S. Versioning in multi-disciplinary engineering with the AutomationML Hub. 4th AutomationML User Conference, Esslingen, Germany, 2016 (upcoming).

[50] Steyskal S., Wimmer M. Leveraging semantic web technologies for consistency management in multi-viewpoint systems engineering. In: Biffl S., Sabou M., editors. Semantic Web for Intelligent Engineering Applications. 1st ed. Springer, New York; 2016.

Chapter 9

The Use of Control Charts in the Study of Bitcoin's Price Variability

Beata Szetela

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66360

Abstract

The focus of this research is bitcoin's variability and its comparison with the variability of the EURO/USD exchange rate. Virtual currencies have been evolving in a dynamic way in the last few years. Under 600 different virtual currencies, the most successful was bitcoin. Its adherents saw in it an alternative to the traditional means of payments allowing the performance of real-time transactions at low costs. The accessibility, where no financial infrastructure is ensured or where either limited or no international agreements exist between financial and banking institutions was also an advantage. The opponents perceived this as a temporary curiosity with no future. Time confirmed that bitcoin has gained on popularity and the exchange rate to the main currencies rose in a dynamic way. The analysts, however, underline that the bitcoin is too volatile and unpredictable, so it cannot compete against the main currencies. The aim of this research is to compare the bitcoin (BTC) to US Dollar (USD) exchange rate and Euro to USD exchange rate volatility using control charts. The results have shown that BTC/USD exchange rate volatility is strongly affected by unexpected price jumps during the period (2010–2016), an act that significantly distinguishes it from more stable and predictable EUR/USD exchange rate variability.

Keywords: bitcoin, virtual currency, control chart, volatility, exchange rate, BTC/USD, EUR/USD

1. Introduction

Bitcoin is a virtual currency and a quite new phenomenon. It was created in 2008 by Satoshi Nakamoto, who published an article "Bitcoin: A Peer-to-Peer Electronic Cash System" in 2008, in which he described a concept of virtual, decentralized and independent means of payment, which is based on a cryptographic blockchain protocol [1]. His main idea was to build a currency based not on trust but on an algorithm, which cannot be influenced or manipulated. It was thought to be independent of any legal or governmental body. After the first release of

bitcoin, by generating the 'block genesis' in 2009, this currency was gaining rapidly in popularity.

As in September 2015 ca. 667 crypto currencies were established.[1] Among them, bitcoin is considered to be the most popular and the most widely used. According to data published on bitinfocharts.com, bitcoin has the highest market capitalization of more than $7 billion, representing 89% of the total capitalization of all cryptocurrency. Further down are the Ethereum $574 million (7.1%), litecoin $179 million (2.2%) and Dash $39 million (0.5%). Data published by blockchain.info at the beginning of 2012 indicated that bitcoin had ca. 400 users. This number increased by ca. 970 k at the beginning of 2014 and reached ca. 8.5 m in September 2016 [2].

Within 4 years since its creation, the European Central Bank (ECB) and other financial institutions still have not come to a final conclusion, as to the classification of bitcoin and other cryptocurrencies. They have not recognized virtual currency as money or as a commodity. As such, the current legislation does not regulate events resulting thereof. The official definition of a 'virtual currency' was set up by a European Central Bank in 2012 for the first time, according to which 'a virtual currency is a type of unregulated, digital money, which is issued and usually controlled by its developers and used and accepted among the members of a specific virtual community' [3]. This definition associated virtual currency with a virtual world in a strict sense and the connection to the real economy was nearly neglected. Nowadays, bitcoin is similar to other virtual currencies that can be traded or exchanged for real money and goods and is accepted by many merchants all over the world. That is why in 2016, the International Monetary Fund (IMF) has extended the definition specifying, that 'virtual currencies can be obtained, stored, accessed and transacted electronically and can be used for a variety of purposes, as long as the transacting parties agree to use them' [4].

According to the IMF, the impact of virtual currencies on the real economy and the financial system is limited. It is, however, possible that with the increase of trading volume and acceptability, virtual currencies can become a serious threat to the financial and banking sector. One of the obstacles to the development of bitcoin named by IMF is its unstable variability.

The purpose of this study is to assess the validity of whether the dollar (USD) is more stable than the bitcoin (BTC) using xbar-s and CUSUM control charts. Depending on the outcome thereof, it will evaluate the concerns of sceptics pertaining to bitcoin and its further development and acceptance in the long term. If it turns out that the volatility of bitcoin does not deviate significantly from the volatility of dollar, then the fears of some financial institutions might appear to be valid.

2. Literature review

Bitcoin is gaining much more popularity not only among financiers but also among scientists. However, a limited number of scientific papers pertaining to cryptocurrency have been

[1] https://bitinfocharts.com/pl/index_v.html

http://dx.doi.org/10.5772/66360

published. Given the above, the proposed content will be an important contribution to studies of both control cards and bitcoin. Current scientific achievements can be divided into four main fields of interest.

- A general and theoretical background concerning the origin, formation and characteristics of bitcoin, e.g. [5–9].
- A number of reports focusing mainly on issues relating to acquisition (mining), trade and broadly understood security, e.g. [10–15].
- The third large group of articles concerns the regulatory environment, including tax-specific policies and possible solutions that regulate the functioning of cryptocurrency in the financial area, e.g. [16–18]. This group also includes various types of reports, publications or banks, financial institutions and government statements on bitcoin, i.e. European Central Bank [19–21] Congressional Research Service acting on the needs of the US Congress [22, 23], Canadian Central Bank [24–26].
- The last group of papers focuses on the application of quantitative methods in the study of Bitcoin.

Generalized autoregressive conditional heteroscedasticity models (GARCHs) were applied to investigate similarities between bitcoin and both US dollar and gold [27, 28]. It was found that bitcoin-like other cryptocurrency tend to generate bubbles and that they do not have fundamental value [29]. Moreover, bitcoin fluctuations are characterized by sudden jumps and extreme pricing, which is characteristic for immature markets [30]. Autoregressive moving average and log-periodic power law models were applied to show that the price of bitcoin depends on the Chicago Board Options Exchange Index Volatility Index, which is indicative of speculation potential [31]. Other scientists argue that fluctuations in the price of Bitcoin are positively correlated with the amount of BTC users and are determined by the shocks of unknown sources of origin. The latter have an endogenous character and are not generated by the impact of specific variables, such as indexes S&P 500, gold rate against the US dollar or (XAU) and the Shanghai stock exchange index (SSE) [32]. According to Bouoiyour et al. [33], bitcoin's price fluctuations are best characterized by a generalized hyperbolic distribution.

The variability of bitcoin against the dollar in 2015 significantly decreased compared with the preceding period. The authors also claimed that bitcoin can be characterized by excessive asymmetry and the price is prone to the negative shocks negative than positive once.

3. Methodology

Statistical process control (SPC) has found its application in many scientific areas. One of the tools, which are used by the SPC, is control charts (see **Figure 1**). A control chart (CC) is a graphical representation of a process. It presents an average value of the quality characteristics reflected on the chart by a central line (CL). Auxiliary lines, called upper control limit (UCL) and lower control limit (LCL), are used for the presentation of deviations from the mean of the process. Control limits are usually set as three times the standard deviation (3-sigma (σ) limits).

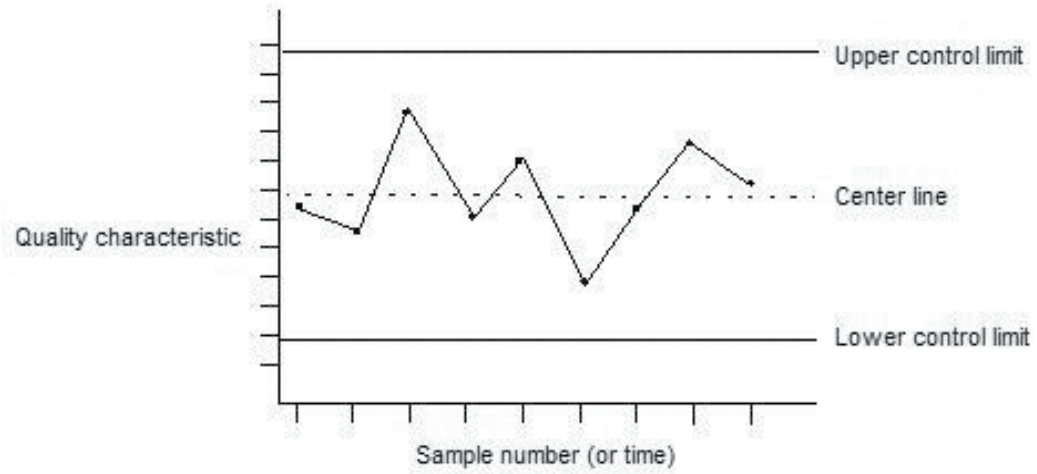

Figure 1. Control chart scheme.

The main idea behind control charts is to monitor an underlying process. If the observations fluctuate in a natural way within set advance control limits and if they do not reveal any specific patterns, then it is said that the process is under control. If, however, the monitored process breaks the established control limits, then it is understood that the process is out of control and specific actions should be launched to return the process under control. For further details considering control charts, please refer to Ref. [34].

SPC uses different types of charts, depending on the type of data used. For the continuous data, the following charts dedicated to variables are used:

- xbar-s charts (controlling the mean and standard deviation of a process),

- xbar-r chart (calculating the mean and the range of a process),

- charts for moving ranges, etc.

For data based on countcharts for attributes are applied, e.g. np chart, p chart, c chart, u chart, etc.

For the purpose of this research, the following control charts will be applied, namely an xbar-s chart and a CUSUM with the moving range chart.

3.1. Xbar-s chart

Xbar-s charts are used to monitor the variation and mean of the process. If the sample size (n) is not constant and is relatively large ($n > 10$), in such a case xbar-s charts are preferable against xbar-R chart.

For the unknown parameters of the s-bar chart and the variable sample size, the central line is defined as an average standard deviation for all samples (Eq. (1))

http://dx.doi.org/10.5772/66360

$$\bar{s} = \sqrt{\frac{\sum_{i=1}^{m}(n_i-1)s_i^2}{\sum_{i=1}^{m}n_i-m}} \tag{1}$$

where m is the number of samples, n_i is the individual sample size and s_i is an individual value of standard deviation for each sample defined as $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$.

The upper (Eq. (2)) and lower control limit (Eq. (3)), which define the boundaries for the (3σ)-three-sigma control limits are calculated based on the following formulae:

$$UCL = \left(1+\frac{3}{c_4}\sqrt{1-c_4^2}\right)\bar{s} \tag{2}$$

$$LCL = \left(1-\frac{3}{c_4}\sqrt{1-c_4^2}\right)\bar{s} \tag{3}$$

where c_4 is a constant.

Similarly, the control limits are determined for the xbar chart, which accompanies the s-chart. The central line is calculated as an average value of the individual averages (Eq. (4))

$$CL = \bar{\bar{x}} = \frac{\sum_{i=1}^{m}n_i\bar{x}_i}{\sum_{i=1}^{m}n_i} \tag{4}$$

The three-sigma control limit is determined by the upper control limit (Eq. (5)) and the lower control limit (Eq. (6)) in the form:

$$UCL = \bar{\bar{x}} + \frac{3\bar{s}}{c_4\sqrt{n}} \tag{5}$$

$$UCL = \bar{\bar{x}} - \frac{3\bar{s}}{c_4\sqrt{n}} \tag{6}$$

It is assumed that the underlying process is under control if it varies between defined control limits. The breach of any of the control limits points at the process being out of control.

3.2. CUSUM and MR chart

Moving range chart enables to plot the sum of ranges of the adjacent pairs of observation within the investigated period. The central line (see Eq. (7)) is calculated as an average range of k-samples and the control limits (see Eqs. (8) and (9)) are set as m-times deviation from the average process range.

$$\overline{R} = \frac{\sum_{i=1}^{k} R_i}{k} \tag{7}$$

$$LCL = \overline{R} - m d_3 \hat{\sigma} \tag{8}$$

$$LCL = \overline{R} + m d_3 \hat{\sigma} \tag{9}$$

where R_i is a range in sample i, k is the number of samples, m is a multiplier chosen to establish control limits, usually set to 3 and d_3 is a constant and $\hat{\sigma}$ is an estimated variance of a process.

CUSUM control charts plot the cumulative sum of deviation from the assumed target value (see Eq. (10)).

$$C_i = \sum_{j=1}^{i} (\overline{x}_j - \mu_0) \tag{10}$$

where C_i is a cumulative sum, $\overline{x}_j$ is mean of a process in sample j, average and μ_0 is target value.

If the process is under control, then a cumulative sum (C_i) follows a random walk process with mean equals 0 [35]. It is assumed that the process is out of control, if the average values drift from the target value. If the values move in the positive direction, then the upper cumulative sum is written as in Eq. (11).

If, however, they move into the negative values, then the lower cumulative sum is defined as in Eq. (12). Finally, if the process exceeds the decision interval, which is contained between the positive and negative sum, then it is assumed, that the process is out of control:

$$C_I^+ = \max\left(0, \overline{X}_i - (\mu_0 + K) + C_{i-1}^+\right) \tag{11}$$

$$C_I^- = \max\left(0, (\mu_0 + K) - \overline{X}_i + C_{i-1}^-\right) \tag{12}$$

where k is a target value, $C_0^+ = C_0^- = 0$, C_i is cumulative sum for sample i.

4. Results

4.1. BTC/USD

The main goal of this research was to compare the variability of two exchange rates: bitcoin to US Dollar (BTC/USD) and Euro to US Dollar (EUR/USD). The average exchange rates for the quarterly data between 2010 and 2016 will be taken into consideration. The final result of the appliance of the xbar-s chart for BTC/USD is presented in **Figure 2**.

It is visible that the process is out of control and the layout of the chart is strongly affected by the significant volatility exchange rate increase in 2013. In the period before, it was a long run

http://dx.doi.org/10.5772/66360

of very low prices. After the peak in 2013, the level of the exchange rate has never reached a comparable value. Taking the above into consideration, the whole investigated period should be divided into consistent and disjoint periods, i.e. the covering time before the positive price shock, namely the year 2013 and the period after 2013. An attempt to analyse the whole period can be misleading due to the faultily estimated control lines understood as a process average and standard deviation.

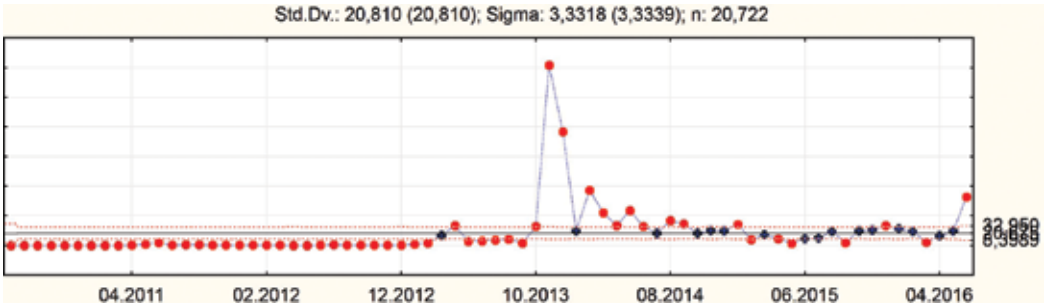

Figure 2. xbar-s chart BTC/USD 2010-2016.

The first studied period was the initial phase of the development of bitcoin. Within that time, the exchange rate has changed significantly, starting from 0.08 USD for 1 BTC, reaching 13.51 USD for 1 BTC at the end of 2012. The xbar-s chart (**Figure 3**) has generated the CL at the level of 0.62. The layout of the chart suggests that the process being out of control in the period reaching April 2011, when the average volatility was significantly under the LCL equalling to 0.31.

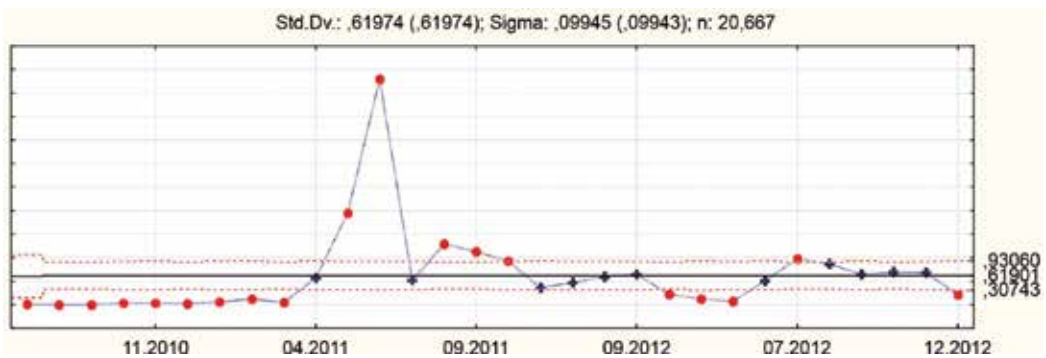

Figure 3. xbar-s chart BTC/USD 2010-2012.

This is justified, as within that time, the price was significantly lower than in other periods. Small changes in prices had almost no effect on volatility. The numbers for May and June have breached the UCL, which equals to 0.93. This peak was connected with strong price jumps, up to 18.50 USD for BTC. The other periods were relatively stable.

A permanent breach of control limits on both sides of CL may be observed, but the deviations are not essential. It is worth to mention that between the years 2010–2012, the graph may imply

a pattern. The observed standard deviations tend to move in the same direction almost every 4 months. The extension to the bitcoin volatility analysis gives the MR chart (**Figure 4**) and CUSUM chart for the standard deviation (**Figure 5**). At the beginning of the investigated process, the exchange rate volatility was dropping constantly. Commencing in April 2011 it is starting to grow, in par with the price increase. After a strong peak in price, which has also a solid and positive effect on price volatility, the bitcoin faced a volatility decrease at the end of 2011, after which the process started to normalize. This phase lasted until the next shock at the end of 2012, where the volatility has started to grow again.

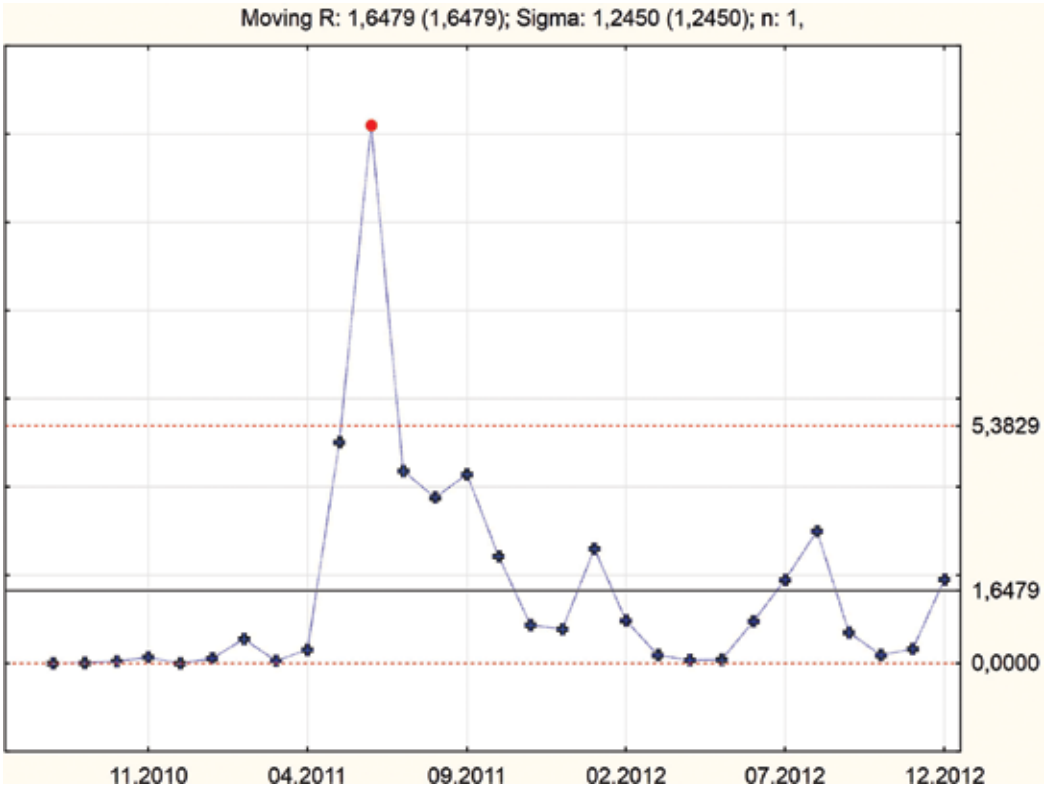

Figure 4. MR chart BTC/USD 2010-2012.

Figure 5. CUSUM BTC/USD 2010-2012.

http://dx.doi.org/10.5772/66360

In 2013, the BTC/USD sudden price jump was observed, which has affected the average price volatility strongly. Throughout the year 2013, the exchange rate was developing steadily (see **Figure 6**). The jump took place between October and November. In this period, exchange rate has risen from 155 to 870. The average price volatility for the year 2013 equals to 51.87. If only the first 10 months of the year were considered, then the CL would be at the level of 12.65.

Figure 6. xbar-s chart BTC/USD 2013.

Moreover, the MR chart (**Figure 7**) shows that within the period between January and October 2013 ranges have fluctuated at the zero line until the exchange rate jumps, which caused an increase of almost 150, thereby, the whole process is said to be significantly out of balance.

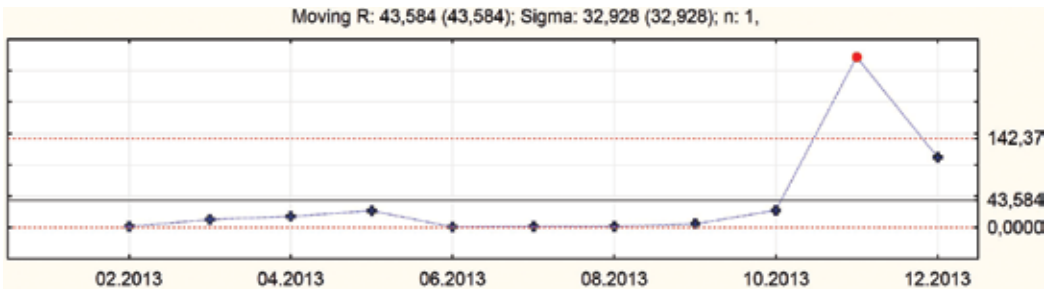

Figure 7. MR chart BTC/USD 2013.

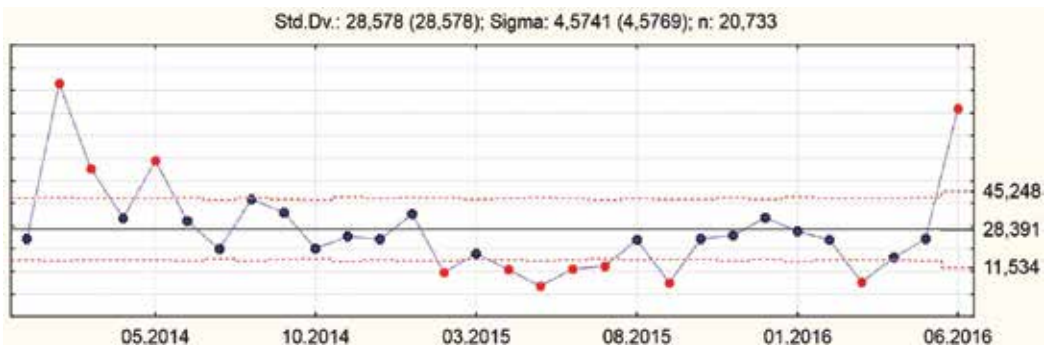

Figure 8. xbar-s chart BTC/USD 2014-2016.

Bitcoin development during the period (2014–2016) was more stable compared to the previous years. The value for the central line in s-chart (**Figure 8**) has declined to 28.4 compared with the year 2013. At the beginning and at the end of the investigated period the price fluctuations

were observed. In January 2014, the exchange rate has reached the level of 649, which was the continuation of the price increase from the year 2013 and in 2016, when the price reached again the limit of 687 US Dollars for a Bitcoin.

The s-chart has produced warning signals concerning a high volatility between February and May 2014 and for June 2016 by breaching the UCL (45.25), as well as signals for low volatility between March and September 2015 by breaching the LCL (11.53). This downturn shift in price was also reflected by the CUSUM chart (**Figure 9**).

These have caused a shift in a process, which was followed by the constant movement with the decreasing tendency of deviations from the process mean until the end of the considered period.

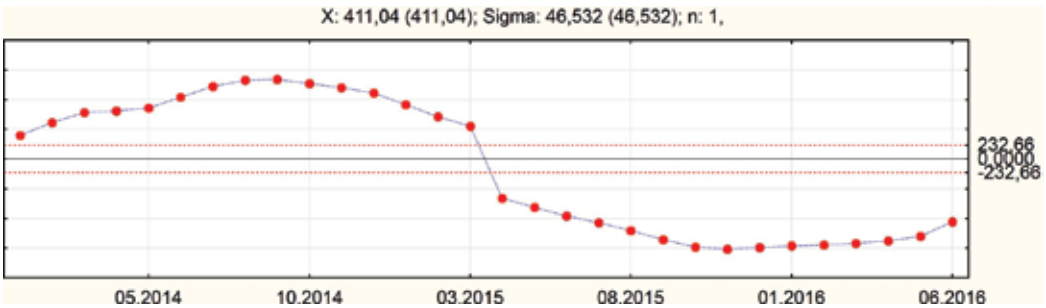

Figure 9. CUSUM chart BTC/USD 2014-2016.

4.2. EUR/USD

The USD/EUR exchange rate volatility seems to be more stable within whole investigated period (see **Figure 10**) compared to the results for BTC/USD (**Figure 2**). The CL generated by the s-chart was at the level of 0.012. The volatility during the years 2010 and 2011 passes the UCL. Global financial crisis has impacted EUR/USD exchange rates strongly. After the crisis hit Greece in 2009 it has moved on and inadvertently affected other European countries, e.g. Spain and Italy. This was the most significant and long-lasting process disruption signalled by the control chart. The other breaches, which appeared in 2014 and at the beginning of 2015, are very close to the LCL (= 0.0052) and UCL (= 0.02). The deviation from CL seems to be not substantial, especially considering the process standard deviation at the level of 0.002.

This development was also reflected by the CUSUM chart (**Figure 11**), where mainly the above average values for the standard deviation were signalled in the year 2010 and the beginning of the year 2011. The rest of the process despite the visible downturn trend remained between the UCL (0.029) and LCL (−0.029).

The central line, which reflects the average standard deviation for the entire process equals 0.1272. If the time after 2013 was considered, then the average standard deviation would be equal to 0.1040, hence both results were at a comparable level.

The EUR/USD exchange rate volatility starting from June 2010 was in the downward trend, which ended in December 2014. The s-chart has produced warning signals for points beyond

http://dx.doi.org/10.5772/66360

the control limits, which were set at a level of 0.02027 for the UCL and 0.00516 for the LCL, respectively. The first signals were produced for the period between August 2010 and October 2011. Strong fluctuations in this time were caused by weak economic data mainly from the United States. It is worth to mention that this was the period short after the economic crisis, when most of the world economies were unstable.

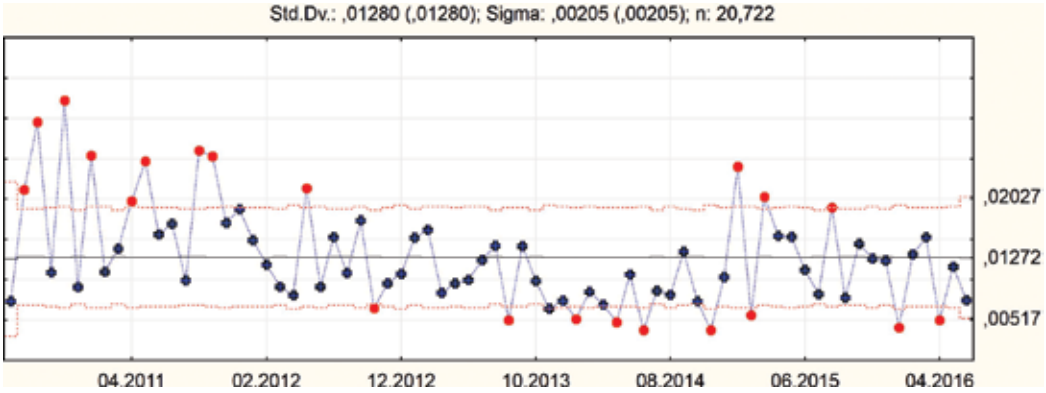

Figure 10. xbar-s chart EUR/USD 2010-2016.

Figure 11. CUSUM chart EUR/USD 2010-2016.

5. Conclusion and further work

Bitcoin, a virtual currency, seems to be a promising alternative to a traditional means of payment. According to a survey published by the IMF, bitcoin has many advantages like low transition costs. It offers the possibility to make transactions with countries with weak financial infrastructure and might contribute to transferring developed technologies and solutions to undeveloped countries. At this point, it is worth mentioning that the idea on which the bitcoin is based has a huge potential in many areas, such as banking, accounting, data gathering and transfer, etc. At the same time, it is not without flaws. Reports and surveys concerning bitcoin mention money laundering issues, low recognisability and lack of stability.

In this chapter, an attempt was made to compare the exchange rate volatility between EUR/USD and BTC/USD. The analysis has shown that the EUR/USD exchange rate volatility is

much more stable compared to the BTC/USD exchange rate. In the entire investigated period, the average exchange rate volatility was at the level of 0.013. The first period after the world crisis, when the international economy was unstable and USA has published economic reports below market expectations, the exchange rate volatility has recorded an increase and charts have produced warning signals for the process being out-of-control. In the remaining period, the exchange rate volatility development did not behave in an unpredictable pattern. The downward trend is visible, but no significant shocks were observed. The lack of sudden fluctuations characterises mature economies.

BTC/USD exchange rate volatility is developing in a completely different way. **Figure 2** shows three establishing phases of this virtual currency; the first one before 2013, when the price level and the overall recognition were nearly zero. The average volatility was equal to 0.6. This value was affected by the price increase from June 2011. In the year 2013, another strong price jump was visible, when the exchange rate has risen from 155 to 870. The third period under consideration was also susceptible to price shocks, namely at the beginning of 2014 and in June of 2016. This unexpected price increase has strongly affected the average exchange rate volatility of BTC/USD. What is positive, the declining tendency in volatility can observe.

In 2013, the average process volatility including outliers caused by price jumps was equal to 51.8, between 2014 and 2016 it has decreased to 28.4. The xbar-s chart and MR-chart for BTC/USD have showed that except for the above-mentioned price fluctuations the process was most of the time under control. It is worth noting at this point that the range between the upper and lower control limits for the bitcoin is broad, which was caused by the extreme price movements. Because of this lack of stability, it is difficult to model bitcoin behaviour. As such, an attempt to forecast its future behaviour based on its past values would be impossible.

The lack of BTC/USD exchange rate predictability and its excessive volatility causes that at least at this stage of development bitcoin cannot threaten the traditional and regarded as stable, currencies such as USD or EUR. It is also associated with a small recognisability and is still limited to a number of places where it can be exchanged or traded. The bitcoin's founder Satoshi Nakamoto said that he is not sure if the bitcoin in this form will survive, but he is convinced that virtual currencies will exist in the future, in this form or in a different one, due to the increasing loss of confidence and trust among the business partners [36].

Author details

Beata Szetela

Address all correspondence to: b.rebisz@gmx.de

Ignacy Łukasiewicz Rzeszow University of Technology, Rzeszów, Poland

http://dx.doi.org/10.5772/66360

References

[1] Blockchain Waller Users. Available from: https://blockchain.info/charts/my-wallet-n-users?timespan=all# [Accessed: 2016.09.05]

[2] Nakamoto, S. http://nakamotoinstitute.org [Internet].2008. Available from: http://nakamotoinstitute.org/bitcoin.pdf [Accessed: 2016.08.09]

[3] European Central Bank. https://www.ecb.europa.eu [Internet]. October 2012. Available from: https://www.ecb.europa.eu/pub/pdf/other/virtualcurrencyschemes201210en.pdf [Accessed: 2016.08.09]

[4] Dong, He, Habermeier, K. F., Leckow,R. B., Haksar, V., Almeida,Y., Kashima, M., Kyriakos-Saad, N., Oura, H., Sedik, T. S., Stetsenko, N., Verdugo-Yepes,C. Virtual Currencies and Beyond: Initial Considerations . IMF Staff Discussion Note. 2016; 3.

[5] Dwyer, G., The economics of bitcoin and similar private digital currencies. *Journal of Financial Stability*, 2015, 17, pp. 81–91.

[6] Dopierała, Ł., Borodo, A. Meaning of the Bitcoin cryptographic currency as a medium of exchange. *Contemporary Economy Electronic Scientific Journal*, 2014, 5(2), pp. 1–12.

[7] Liu, J., Kauffman, R., Ma, D. Competition, cooperation and regulation: Understanding the evolution of the mobile payments technology ecosystem. *Electronic Commerce Research and Applications*, 2015, 14(5), pp. 372–391.

[8] Jagwani, B. Bitcoins demystified. *International Journal of Marketing, Financial Services & Management Research*, 2015, 4(4), pp. 29–35.

[9] Rogojanu, A., Badea, L. The issue of "true" money in front of the bitcoin's offensive. *Theoretical and Applied Economics*, 2015, 22(2(603)),pp. 77–90.

[10] Badev, A., Chen, M. Bitcoin: Technical Background and Data Analysis. Finance and Economics Discussion Serie, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C., 2014.

[11] Luther, W. J., Olson, J. Bitcoin is memory. *Journal of Prices & Markets*, 2015, 3(3), pp. 22–33.

[12] Campbell, H. R. [Updated: 2016.08.17]. 2014. Available from: http://ssrn.com/abstract=2479670

[13] Tasca, P. i De Roure, C. Bitcoin and the PPP Puzzle. [Updated: 2016.08.17]. 2014. Available from: http://ssrn.com/abstract=2461588

[14] Karame, G. O. androulaki, E., Roeschlin, M., Gervais, A., Čapkun, S. Misbehavior in bitcoin: A study of double-spending and accountability. *ACM Transactions on Information and System Security*, 2015, 18(1). Article No.: 2

[15] Mandjee, T. Bitcoin, its legal classification and its regulatory framework. *Journal of Business & Securities Law*, 2015, Vol. 15 (2). p.157–217.

[16] Bryans, D. Bitcoin and money laundering: Mining for an effective solution. *Indiana Law Journal*, 2014, 89(1). [Updated: 2016.08.17]. Available from: http://www.repository.law.indiana.edu/ilj/vol89/iss1/13

[17] Plassaras, N. A. Regulating digital currencies: Bringing bitcoin within the reach of IMF. *Chicago Journal of International Law*, 2013, 14, 377–408.

[18] ECB. *Virtual Currency Schemes*. [Updated: 2016.08.17]. 2012. Available from: https://www.ecb.europa.eu/pub/pdf/other/virtualcurrencyschemes201210en.pdf

[19] ECB. Virtual Currency Schemes—A Further Analysis. Frankfurt am Main: European Central Bank. [Updated: 2016.08.17]. 2015. Available from: https://www.ecb.europa.eu/pub/pdf/other/virtualcurrencyschemesen.pdf abgerufen

[20] Draghi, M. A letter from the President of the ECB to Mr Buonanno, the Honourable Member of the European Parliament, regarding the relevance of virtual currency schemes (VCS) for the ECB's tasks and their potential impact on policies in these specific areas. [Updated: 2016.08.17]. 2015. Available from: http://www.ecb.europa.eu/pub/pdf/other/150421letter_buonanno_3.en.pdf

[21] Murphy, E. V., Murphy, M., Seitzinger, M. Bitcoin: Questions, Answers and Analysis of Legal Issues. [Updated: 2016.08.17]. 2015. Available from: https://www.fas.org/sgp/crs/misc/R43339.pdf

[22] Congress, T. L. Regulation of Bitcoin in Selected Jurisdictions. Global Legal Research Directorate Staff. The Law Library of Congress. [Updated: 2016.08.17]. 2014. Available from: http://www.loc.gov/law/help/bitcoin-survey/regulation-of-bitcoin.pdf

[23] Gans, J. S., Halaburda, H. Some Economics of Private Digital Currency. Bank of Canada Working Paper(38). [Updated: 2016.08.17]. 2013. Available from: http://www.bankofcanada.ca/wp-content/uploads/2013/11/wp2013-38.pdf

[24] Bank of Canada. Decentralized E-Money (Bitcoin) [Internet]. 2014. Available from: http://www.bankofcanada.ca/wp-content/uploads/2014/04/Decentralize-E-Money.pdf [Accessed: 2016.08.17]

[25] Chiu, J., Wong, T.N. On the essentiality of E-money. Staff Working Paper(43). 2015. Available from: http://www.bankofcanada.ca/wp-content/uploads/2015/11/wp2015-43.pdf [Accessed: 2016.08.17]

[26] Dyhrberg, A.H. Hedging capabilities of bitcoin. Is it the virtual gold? *Finance Research Letters*, 2016, Vol. 16. pp. 139–144. .

[27] Dyhrberg, A. H. Bitcoin, gold and the dollar—A GARCH volatility analysis. *Finance Research Letters*, 2016, Vol. 16. pp. 85–92.

[28] Cheah, E.T., Fry, J. Speculative bubbles in bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, 2015, 130, pp. 32–36.

http://dx.doi.org/10.5772/66360

[29] Gronwald, M. The Economics of Bitcoins—Market Characteristics and Price Jumps. CESifo Area Conference on Macro, Money and International Finance. Munich. 20–21.02.2015

[30] MacDonell, A. Popping the Bitcoin Bubble: An application of log-periodic power law modelling to digital currency. University of Notre Dame Working Paper, 2014.

[31] Vockathaler, B. The Bitcoin Boom: An in Depth Analysis of the Price of Bitcoins. Major Research Paper University of Ottawa. 2015.

[32] Chu, J., Nadarajah, S., Chan, S. Statistical analysis of the exchange rate of Bitcoin. *PLoS One*. 2015, 10(7), e0133678.

[33] Bouoiyour, J., Selmi, R. Bitcoin Price: Is it really that New Round of Volatility can be on way? *MPRA Paper No. 65580*. 2015.

[34] Douglas, C. MontgomeryNakamoto Montgomery. Introduction to Statistical Quality Control. 7th ed. Wiley, USA. 2013.

[35] Harris, T. J., Ross, W. H. Statistical process control procedures for correlated observations. *Canadian Journal of Chemical Engineering*, 1991, 69(1), pp. 48–57.

[36] Re: Nakamoto, S. Bitcoin v0.1 released. 2009. Available from http://satoshi.nakamotoinstitute.org/emails/31 cryptography/17/ [Accessed: 2016.09.05]

Edited by Leo D. Kounis

Quality control and assurance cover a diverse area of modern life and play, undeniably, an important role. This book brings together a collection of international papers that showcase examples of current research and practice in industry and the medical profession. It is hoped that engineers, researchers and scientists will be assisted in their continuous quest for excelling in qualitative aspects. The Ancient Greek word arete means excellence or virtue and defines the highest qualitative state: a man’s effectiveness and skill in goodness (optimum potentiae). Indeed, Ancient Greeks believed that without quality control, specifications are useless and may result to illegitimacy, which in turn may become a threat to society itself.

ISBN 978-953-51-6696-2

Photo by Milkas / iStock

IntechOpen