Towards a maturity model for the assessment of data management of healthcare entities in developing countries

by

Laubscher van der Merwe

Thesis presented in fulfilment of the requirements for the degree of

Master of Engineering (Industrial)

in the Faculty of Engineering at Stellenbosch University

Supervisor: Dr IH de Kock
Co-supervisor: Dr WG Bam

March 2021

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 19/11/2020

Abstract

The study focused on healthcare data management in developing countries. Data management is important for good healthcare service delivery. Data helps management with decision-making based on facts and helps care delivery to patients considering historical data. Health indicators can also be collected for population health surveillance. But for data to be of beneficial use in the health sector, it needs to be accurate, consistent, available and secured. The health sector in developing countries faces many data management challenges and struggles to provide patients with efficient and cost-effective care.

These data management challenges span the whole scope of the healthcare data value chain. In healthcare, there are ample challenges with: (i) data integration; (ii) human factors; (iii) data collection; (iv) data security; (v) data quality; (vi) infrastructure and technology; (vii) data transmission; (viii) the implementation of systems; (ix) data retrieval; and (x) data governance.

Improving the management of healthcare data can lead to improved healthcare service delivery. To improve healthcare data management effectively, it is important to determine which components to focus on. Therefore, the aim of the study was to develop a maturity model that assesses the as-is state of a healthcare delivery entity's data management across the whole healthcare data value chain, which then identifies the areas of improvement.

Maturity models are designed to assess the maturity of a selected domain based on a set of criteria. They are artefacts that determine the *status quo* of the capabilities of an organisation and derive measures to improve from there. The basic purpose of maturity models is to outline the maturity levels that can be used to make maturity assessments. This includes the description of the characteristics of each level and the logical relationship between levels.

Therefore, this study investigated the use of a maturity model as a suitable research product to assess the healthcare data management in developing countries. To this end, a customised tool called the Healthcare Data Management Maturity Model (HCDMMM) was developed to assess data management of healthcare delivery entities in developing countries, including hospitals and clinics at the facility level, and the headquarters at the organisational level. Both facility and organisational levels were included in the study to address data management challenges across the whole data management system in healthcare.

Throughout the development process of the HCDMMM, the structure and components were verified for their theoretical soundness. After the development process, it was further verified against the specified requirements for the proposed solution. The HCDMMM was also validated to ascertain whether it is useful to its intended users by evaluating it against the dimensions of applicability, practicability, usability, and determining strengths and weaknesses.

Opsomming

Die studie handel oor gesondheidsorgdatabestuur in ontwikkelende lande wat vir goeie gesondheidsorgdienslewering belangrik is. Data help bestuurslede met feitegegronde besluitneming en ondersteun die voorsiening van gesondheidsorg aan die hand van historiese tendense. Gesondheidsaanwysers kan ook ingesamel word om populasiegesondheid te moniteer. Data kan egter slegs van nut wees vir die gesondheidsektor indien dit akkuraat, konsekwent, beskikbaar en beskerm is. Die gesondheidsektor in ontwikkelende lande word deur etlike databestuursuitdagings in die gesig gestaar, en sukkel om doeltreffende en koste-effektiewe sorg aan pasiënte te voorsien.

Hierdie databestuursuitdagings strek oor die volle spektrum van die gesondheidsorgdatawaardeketting. Op gesondheidsorggebied is daar heelwat uitdagings met (i) data-integrasie, (ii) menslike faktore, (iii) data-insameling, (iv) datasekuriteit, (v) datagehalte, (vi) infrastruktuur en tegnologie, (vii) dataoordrag, (viii) stelselsimplementering, (ix) dataherwinning, en (x) databestuur.

’n Verbetering in gesondheidsorgdatabestuur kan ook beter gesondheidsorgdienslewering teweegbring. Om gesondheidsorgdatabestuur doeltreffend te verbeter, is dit belangrik om te bepaal presies watter komponente aandag moet ontvang. Vir dié doel ontwikkel hierdie studie ’n volwassenheidsmodel wat die huidige stand van ’n gesondheidsorgentiteit se databestuur oor die hele gesondheidsorgdatawaardeketting beoordeel, wat dan gebiede vir verbetering aan die lig bring.

Volwassenheidsmodelle word ontwerp om die volwassenheid van ’n bepaalde domein op grond van ’n stel kriteria te bepaal. Dié modelle is instrumente wat die status quo van ’n organisasie se vermoëns bepaal en op grond dáárvan dan maatreëls vir verbetering identifiseer. Die hoofdoel van volwassenheidsmodelle is om deur volwassenheidsvlakke volwassenheidsbeoordelings te onderneem. Dít sluit in die beskrywing van die eienskappe van elke vlak en die logiese verwantskap tussen vlakke.

Hierdie studie ondersoek dus die gebruik van ’n volwassenheidsmodel as ’n geskikte navorsingsproduk om gesondheidsorgdatabestuur in ontwikkelende lande te beoordeel. Hiervoor word ’n pasgemaakte instrument genaamd die Volwassenheidsmodel vir Gesondheidsorgdatabestuur (oftewel “Healthcare Data Management Maturity Model” (HCDMMM)) ontwikkel om die databestuur van gesondheidsorgentiteite in ontwikkelende lande, waaronder hospitale en

klinieke op fasiliteitsvlak, en hoofkantore op organisasievlak, te evalueer. Sowel die fasiliteits- as organisasievlak word by die studie ingesluit om die databestuursuitdagings oor die hele databestuurstelsel van gesondheidsorg die hoof te bied.

Die struktuur en komponente van die HCDMMM is deur die hele ontwikkelingsproses vir teoretiese korrektheid nagegaan. Boonop is dit ná die ontwikkelingsproses verder gestaaf aan die hand van die bepaalde vereistes waaraan die voorgestelde oplossing moet voldoen. Daarbenewens is die nut van die HCDMMM vir die beoogde gebruiker bevestig deur die toepaslikheid, uitvoerbaarheid en bruikbaarheid daarvan te evalueer en die sterkpunte sowel as swakpunte van die model te bepaal.

Acknowledgements

Firstly, I want to extend my sincere gratitude towards my supervisors, Ms Imke de Kock and Dr Wouter Bam. Both of you played major roles in my research journey and supported me in different ways.

I am also thankful towards all the subject matter experts who offered their time to help me develop, verify and validate my research. Their contributions were invaluable.

I want to thank all my faithful friends who were a part of this journey with me, who encouraged me through the bad times and who supported me when I need it. Thank you for all the great times that took my mind off my studies to keep me sane.

To my parents, Willem and Ronelle, I am forever grateful for your continuing care and support. Thank you that I can always count on you for anything.

I want to thank every individual who has invested in me and taught me so much over the course of these two years. Thank you for enriching my life and empowering me with skills, knowledge, wisdom and love.

Lastly, I also want to extend my thanks to my siblings, Martina, Ansa and Roelof Willem, who is and will be my life-long friends. Thank you for everything we have been through and thank you for all your support.

Contents

List of figures

List of tables

Nomenclature

Glossary

Domain component: Domain components are significant particular parts of a specific domain that plays a major role in the maturity of the domain.

Domain sub-components: It is the distinct capability areas within the domain components that allow for further detail of the domain components. The capability area relates to the ability of a system to perform a specific function, process or cluster of activities.

Domain: The sphere of activity that defines the broad focus of the entity under study.

Enabling practices: The enabling practices are the conventions that shape the working of the primary activities. These are customs that creates the environment that the primary activities are carried out in.

Facility level: The facility level is the micro-level of the healthcare data management system where micro-functions are carried out. This is the operational level and functions that are carried out has to do with operational activities.

Maturation path: The anticipated, desired or logical path a capability area undergoes towards a target maturity state from an immature state.

Mature state: The state where the capability area and the domain under study is continuously improved in terms of effectiveness and efficiency.

Maturity model: A maturity model is designed to assess the maturity of a selected domain based on a set of criteria. They are artefacts that determine the status qua of the capabilities of an entity and derives measures to improve from there. A maturity model consists out of stages for each capability area that form an anticipated, desired or logical path from an initial to a target maturity state.

Maturity: The extent of how well a specific function or process is performed.

Organisational level: The organisational level is the meso- and macro-level of the healthcare data management system where the meso- and macro-functions are carried out. This is the managerial level and functions on this level often affects functions on the micro-level too. The functions on this level has an organisation-wide effect.

Primary activities: The primary activities are the core activities in every domain component to accomplish the goal of that specific domain component.

Supporting structures: Supporting structures are the secondary capability areas in the domain components that support the working of the core activities to accomplish the domain component goal. These capability areas are the specific secondary inputs and services needed to accomplish the goals of the primary activities.

System levels: Different system levels describe the different levels of the system that different functions are carried out on. The functions of the different levels are addressed to different audiences of the system.

Acronyms

AP Attention Point

BC Boundary Condition

CA Capability Area

CMM Capability Maturity Model

CMMI Capability Maturity Model Integration

DA Data Analytics

DR Design Restriction

DS Design Science

DSR Design Science Research

EHR Electronic Health Records

EMR Electronic Medical Record

EP Enabling Practices

FR Functional Requirement

GDPR General Data Protection Regulation

HAAM Healthcare Analytics Adoption Model

HCDMMM Healthcare Data Management Maturity Model

HCW Healthcare Worker

HIE Health Information Exchange

HISAM Healthcare Information Security Adoption Model

HISMM Hospital Information System Maturity Model

HR Human Resources

ICMM Informatics Capability Maturity Model

ICT Information and Communications Technology

IoT Internet of Things

IT Information Technology

KPI Key Performance Indicator

MM Maturity Model

MMEI Maturity Model of Enterprise Interoperability

MVoT Multiple Versions of Truth

NHS National Health Service

PA Primary Activity

POPIA Protection of Personal Information Act

RDMS Relational Database Management Systems

SME Subject Matter Expert

SQL Structured Query Language

SS Support Structure

SSoT Single Source of Truth

UR User Requirement

WHO World Health Organization

Chapter 1

Introduction

This chapter introduces the research and gives background on healthcare data management and its challenges in developing countries. The background on healthcare data management in developing countries that is described in Section 1.1 includes the importance of data management for effective care delivery, basic data management components, some common challenges of healthcare data management and why it is important to improve healthcare data management in developing countries. This chapter also contains the research problem statement and the research aim and objectives (Section 1.2). The research strategy is also specified (Section 1.3), the scope of the research is set (Section 1.4) and the document structure is conveyed (Section 1.5). Lastly, the chapter is concluded (Section 1.6).

1.1 Background

Data management is a significant contributor to the effectiveness of any healthcare system. The objective of health data management is to produce relevant and quality data to support health interventions (World Health Organization, 2008). In developing countries, however, many data management challenges exist that impede the effectiveness of the healthcare system.

Data management mainly entails the collection, storage, security and sharing of data gained from diverse sources (Galetto, 2016). Data management in healthcare is the basis that enables the holistic views of patients, personalisation of treatments, improved communication and enhancement of health outcomes (Evariant, 2019). To achieve this, data must be aggregated and standardised (Adams, 2017). The accuracy, completeness and consistency of data must be ensured throughout the healthcare data management system. Therefore, a data management plan is needed, coupled with the necessary platform to integrate data, manage its quality and utilise it productively.

Data is used in all aspects of the healthcare system. Data is collected, stored and used for patient record keeping, monitoring, diagnosis and treatment. It is also used in other parts of the healthcare system such as tracking medicinal stock levels and patient billing. Healthcare data is located everywhere: (i) clinical and claims systems; (ii) human resources (HR); (iii) financial applications; and (iv) third-party sources (Adams, 2017). Health data is also crucial in the case of emergent diseases and other acute health threats (World Health Organization, 2008). The rapid awareness, investigation and response to these health threats can save lives and prevent broader national or global outbreaks.

Some of the common challenges that arise in developing countries such as South Africa are: (i) the fragmented nature of information systems or the lack of integration (Bhaskaran *et al.*, 2013; Masana and Muriithi, 2017); (ii) challenges with data collection (Allorto and Wise, 2015), storage (Ganiga *et al.*, 2018), retrieval (Kaseke *et al.*, 2017) and sharing (Pahl *et al.*, 2015); (iii) data security challenges (Khan and Hoque, 2016*b*); (iv) data quality challenges (Turan and Palvia, 2014); (v) poor governance (Alkraiji *et al.*, 2016); (vi) human incompetence (Mate *et al.*, 2009; Sharifi *et al.*, 2013); (vii) and the lack of the necessary infrastructure for data management (Fritz *et al.*, 2015; Bijlmakers *et al.*, 2017). Data management challenges not only contribute to the ineffective delivery of healthcare, but also hinder the application of novel, innovative solutions in developing countries. For instance, South Africa is developing an 'mhealth' strategy to improve the accessibility of healthcare. Mhealth is dependent on the availability and accessibility of good quality data to have any significant impact (South-African Department of Health, 2015). Other applications that are dependent on the availability of data are data science, the Internet of Things (IoT), cloud computing, telehealth and machine learning. These applications can only be applied in healthcare if the management of data is done properly. Data management makes it possible to handle large volumes of structured or unstructured data. Through data management best practice the power of the data can be harnessed and insights can be gained to make the data useful (Galetto, 2016).

Research in the field of healthcare data management of developing countries is important, because data is such an integral part of the healthcare system. The healthcare system will not be able to function properly without data management. Data management in developing countries is still in its infancy stage, with developing counties generally lacking the needed capabilities to execute healthcare data management strategies. Although data management has been a part of healthcare of developing countries for a long time, developing countries still struggle with elementary challenges and are only starting to take the initial steps to the development of mature capabilities of which developed countries boast.

There are many reasons why it is important to strengthen healthcare data management. There exists an increase in the demand for better statistics that accurately track progress and performance in health and to ensure accountability, on both country and global levels (World Health Organization, 2014). Reliable, regular and timely data is necessary to monitor the progress of different countries who strive to improve service delivery, reach Sustainable Development Goals (SDGs) and Universal Health Coverage (UHC) targets. Better health data and information leads to better decision-making, which results in better health (World Health Organization, 2008). Reliable and timely health data serves as a foundation for public health action and the improvement of health systems on a national and international level. Available public health surveillance data can be used for defining problems and provide timely action plans. These are but a few reasons why it is important to improve healthcare data management.

1.2 Problem statement, research aim and objectives

In this section the problem for this study (Section 1.2.1) and the research aim (Section 1.2.2) are stated. The research aim is further expanded with a set of research objectives (Section 1.2.2).

1.2.1 Problem statement

Healthcare data management is a significant challenge in developing countries. There are many aspects contributing to the sub-optimal state of the healthcare data management system. Effective data management is supposed to improve the service delivery of healthcare, but many factors contribute to the ineffective management of data such as the lack of infrastructure, less qualified staff, paper-based data collection and storage systems, data quality and security challenges and the lack of integration of the data management system restraining data sharing or the dissemination of data to patients. This results in ineffective data management that does not adequately improve the service delivery of healthcare.

Structured tools that assist stakeholders to identify the strengths and weaknesses of their existing data management practices and present a pathway for improvement can greatly assist systematic improvement initiatives and planning. Therefore, a need exists for tools that enable this by (i) clearly identifying the various aspects of data management that need to be considered in the healthcare centre and (ii) integrating these into a practical tool for assessing both the current strengths and weaknesses to assist in determining the pathway for improvement of data management practices within the healthcare sector.

1.2.2 Research aim and objectives

The aim of the study was to contribute towards increasingly effective and efficient healthcare data management practices by developing a tool that enables the identification of the strengths and weaknesses of existing data management practices that assists systematic improvement initiatives and planning. The tool should be able to this by (i) clearly identifying the various aspects of data management that need to be considered in the healthcare centre and (ii) integrating these into a practical tool for assessing both the current strengths and weaknesses of existing data management practices within the healthcare sector to assist in determining the pathway for improvement of these data management practices. The set of research objectives that is developed to collectively contribute towards the realisation of the stated aim is outlined below. The research objectives and sub-objectives are listed below with the related chapters and sections where they are addressed:

1. Describe the context of healthcare and healthcare data management in order to gain a better understanding of healthcare data management in developing countries (Chapter 2). The sub-objectives that relate to research objective 1 include:

    1.1 describing the delivery of healthcare as a system (Section 2.1);

    1.2 defining data management and describing data management in the context of the healthcare sector (Section 2.2); and

    1.3 identifying the significant challenges of healthcare data management in developing countries (Section 2.3).

2. Specify the requirements that the proposed research product should address in order to ensure the research product adequately addresses the problem (Chapter 3). The sub-objectives that relate to research objective 2 include:

    2.1 identifying and describing the significant healthcare data management components (Section 3.1);

    2.2 determining the challenges landscape of data management across the whole healthcare data management value chain (Section 3.1.2); and

    2.3 specifying the requirements that the proposed research product should address to be able to assist in identifying healthcare data management components to improve on and to address the problem statement of this study (Section 3.2).

3. Identify, select and describe a suitable research product that is able to facilitate the identification of healthcare data management components to

improve on in order to address the healthcare data management problem stated for this study (Chapter 4).

4. Develop the research product in order to provide an appropriate means to identify the healthcare data management components to focus on for improvement endeavours (Chapter 5 and 6). The sub-objectives that relate to research objective 4 include:

   4.1 using a defined design and development methodology with the appropriate design decisions and principles to develop the research product (Sections 5);

   4.2 verifying the research product theoretically and whether it addressed the specified requirements (Section 6.2); and

   4.3 validating the applicability, practicability and usability of the developed research product and determine its strengths and weaknesses (Section 6.3).

1.3 Research strategy

There are many challenges associated with the management of healthcare data and it is important to follow an appropriate research strategy to propose a suitable research product. Figure 1.1 illustrates the different components of the research strategy for this study which consists of: (i) a literature study to understand relevant concepts; (ii) a requirements mapping; (iii) the research product development process; (iv) and verification and validation of the research product. The different components of the research strategy are described further:

1. **Literature study to understand relevant concepts**

   Literature was studied to understand the healthcare context, data management concepts, as well as data management, in the context of healthcare. A structured literature review was conducted to determine the scope of healthcare data management challenges in developing countries. This scope of challenges defined the healthcare data management problem holistically and provided a platform on which to develop a framework to address the healthcare data management problem. Literature was further reviewed to gain knowledge on the healthcare data management value chain from a data analysis perspective. This value chain was used to construct the challenges landscape that was used for the requirements mapping of the proposed framework. Literature was further investigated to find maturity models to determine whether they can address the healthcare data management problem in developing countries. The literature of maturity models included the origin and purpose of maturity models, the basic structure of maturity models, and using a defined design methodology that included design decisions and principles. Literature on existing maturity models of healthcare data management was also reviewed to establish what challenges these models addressed in the past, what can be learnt from them and how the developed model for this study is different from previous ones.

2. **Requirements mapping**

   The proposed framework to address the healthcare data management problem should adhere to certain requirements. General requirements that should be met were specified by using the value chain, challenge landscape and other relevant literature as inputs for the requirements specification. The categories of the different design specifications as specified by Van Aken and Berends (2007) were used as a guide for the requirements mapping.

3. **Research product development process**

   The research product design requirements and development procedure gained from literature were then used to develop a research product that meets the design requirements that address the healthcare data management problem in developing countries. An iterative design process was followed that utilised inputs from literature and subject matter experts (SMEs) to develop the proposed research product with all the necessary components.

4. **Verification and validation**

   SMEs were interviewed to ensure the correct design of the model and to ensure it appropriately represents the real-world system. Throughout the study, literature was consulted to verify the research process and SMEs were interviewed to verify the developed components of the model. The model was developed iteratively and improved incrementally with the assistance of SMEs to ensure the model was theoretically sound. The model was validated to ensure it is suitable for use in the real world by real users. It was validated along the dimensions of applicability, practicability and usability to determine whether it attained the aim of the study.

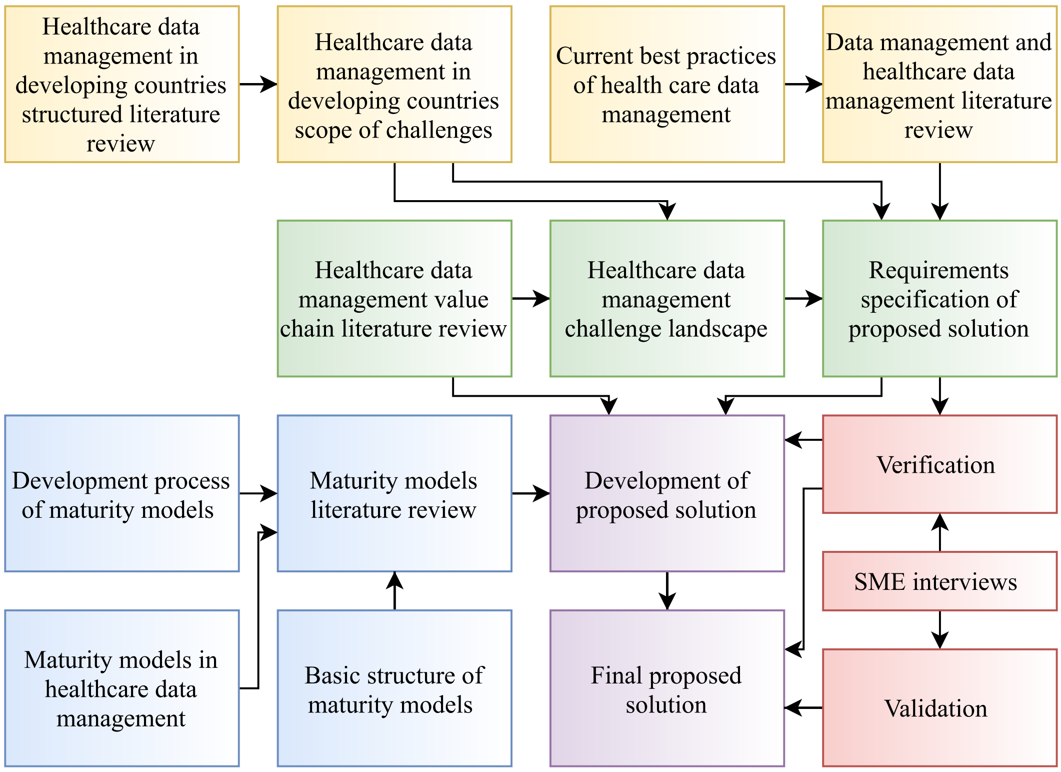


Figure 1.1: Research roadmap

1.4 Scope of the research

In this section the delimitations and limitations of the study are conveyed. The delimitations (Section 1.4.1) and limitations (Section 1.4.2) together comprise the scope of the research and fix the boundary of the study.

1.4.1 Delimitations

This study focuses on healthcare data management in developing countries. It investigated healthcare data management from a system perspective, with a specific focus on the facility and the organisational levels of the healthcare system. Data management on the facility level included data management of hospitals and clinics, and on the organisational level, it included data management of the organisation's headquarters. This study investigated healthcare data management from a strategic level and therefore, did not describe all the detailed components of the healthcare data management system. To establish the different domain components of healthcare data management, a healthcare data value chain was constructed. The study focused on the development of a maturity model in the healthcare data management domain to help assess the as-is state of the capabilities of the technical components of healthcare entities' data management. Although the model development focused on healthcare data management in the public sector of developing countries, it was also validated whether the model is also applicable to the private sector.

1.4.2 Limitations

Healthcare data management in developed countries has a very different focus from that of developing countries, especially compared to data management in the public sector, and therefore, this study is not relevant to healthcare data management in developed countries. This study did not investigate data management from a sociotechnical perspective, but focused only on the technical components of data management in healthcare. Neither does it offer any solution for challenges related to healthcare data management cost and financial support. This study does not focus on identifying and describing the healthcare data management components on a detailed level, but merely describes their functions on a strategic level. The maturity model that was developed is limited to a descriptive application of use and does not provide a means of progressing to the next maturity stages, although its descriptions of levels give an indication of the required capability on each maturity level. Lastly, due to time constraints, this study does not include the application of the developed model in a real-world setting as part of validation.

1.5 Document structure

The presentation of the research of this study guides the reader from understanding the background and context of the challenges that healthcare data management in developing countries faces, to understanding the theoretical foundations that were used to develop the model, whereafter the model that was developed is presented. The evaluation of the model is also described. What follows is a brief description of the contents of each chapter.

- **Chapter 1: Introduction**

  This chapter gives a summary of the context within which the research problem exists. The research problem statement, aim and objectives are also stated. Thereafter, the research strategy is described, the scope of the research is set and lastly, the document structure is discussed.

- **Chapter 2: Contextualisation of data management in healthcare**

  This chapter focuses on understanding data management and its challenges in the context of healthcare in developing countries. This is done by explaining the healthcare system, describing data management and explaining data management specifically in the context of healthcare.

This chapter concludes with the scope of the healthcare data management challenges that were determined through a structured literature review.

- **Chapter 3: Requirements specification**

  In this chapter the requirements that the proposed research product should adhere to are specified. In order to specify these requirements, it was necessary to determine the challenges landscape across the whole healthcare data value chain on which the requirements would be based. The healthcare data value chain was developed by incorporating different proposed big data value chains. To determine the challenges landscape, the challenges from the scope of challenges in Chapter 2 were categorised into the different healthcare data value chain components. The healthcare data value chain and challenges landscape is covered in this chapter. Following this, the specified requirements are described as they were determined from Chapter 2 and 3 inputs.

- **Chapter 4: Maturity models**

  The focus of this chapter is to describe a maturity model as an appropriate research product to help address the specified problem of this study. The origin, purpose and value of maturity models are portrayed, followed by the basic structure of maturity models. The importance of using a defined methodology to develop maturity models is also conveyed. The chapter concludes with the review of maturity models in the healthcare data management domain.

- **Chapter 5: Model development and presentation**

  This chapter focuses on describing the development of the maturity model to help address the healthcare data management challenges in developing countries. Firstly, the development methodology is described, followed by explaining how the development methodology was executed, which included the identification of the need for new opportunity, the definition of the scope of the model, and the iterative design and population phase of the maturity model. Finally the conceptual maturity model is presented, along with the maturity model transfer media (the form in which the maturity model is transferred to user communities). The last part of this chapter described the reflection on future updates of the developed model.

- **Chapter 6: Model evaluation**

  In this chapter the evaluation of the developed maturity model is described. The evaluation strategy, which consists of verification and validation, is described. Thereafter, the execution of the verification process is conveyed. The last section of this chapter is the explanation of how the validation process was executed.

- **Chapter 7: Conclusion**

  This final chapter includes an overview of the research conducted in this study, the confirmation that the stated research objectives have been achieved, a discussion of the limitations of the research, and a proposal of opportunities for future research.

1.6 Chapter conclusion

This chapter contains an introduction of the challenges of healthcare data management in developing countries. It also defines the problem statement, and the research aim and objectives. The research strategy is also explained. Thereafter, the scope of the research is set and the last section of this chapter consists of the description of the document structure. Following this chapter, the context of data management in healthcare in developing countries is described in detail. This provides a detailed exposition of the healthcare data management challenges in developing countries.

Chapter 2

Healthcare data management in developing countries and its scope of challenges

Healthcare data management can considerably improve the outcomes of healthcare systems. It can enable effective patient record-keeping, improve diagnosis, treatment and monitoring. The proper use of data can improve the quality of healthcare, but to achieve this, the appropriate systems and infrastructure need to be in place. Over the years it has been the constant struggle of developing countries to properly manage their data for beneficial use. Data management challenges surface throughout the healthcare system and impede the beneficial use of data. To be able to contribute to improving the healthcare data management problem, it is important to have a basic understanding of the healthcare system and data management in healthcare. It is also important to gain more insight into the persistent healthcare data management challenges in developing countries. Therefore, in this chapter, the healthcare system is described to give background to this domain (Section 2.1). Data management in healthcare is also described to gain an understanding of data management in the context of healthcare (Section 2.2). To determine the persistent healthcare data management challenges the current healthcare data management challenges in developing countries were scoped through the use of a structured literature review and the significant challenges contributing to the healthcare data management problem were identified (Section 2.3). This chapter is concluded in Section 2.4.

2.1 The healthcare system

To understand the healthcare data management problem better, it is important to gain better insights into the healthcare sector itself. This section describes the healthcare sector looking at healthcare as a system. First, the systems approach is described in general terms (Section 2.1.1) and then it is applied to the context of the healthcare sector (Section 2.1.2). The healthcare sector is described as a system with various levels.

2.1.1 The systems approach

The systems approach is what systems engineering is founded on and is a set of top-level rules from which systems engineering methodologies are derived (Jackson *et al.*, 2010). Systems approach is fundamental to systems, systems thinking, systems methodology, systems design and systems engineering.

A system consists of a collection of interacting elements (Jackson *et al.*, 2010). This is called a system of interest. The system of interest can interact with other systems in the environment it operates in. The combination of different systems that interact with each other constitutes a system of systems. To define the system of interest, it is important to define the system boundary which clearly distinguishes the system from its environment. Internal-to-external interactions take place across the system boundary. An open system exchanges energy, information and material across its boundary with its environment and other systems in the environment. Environments of interests can include physical, cultural, economic, social, legal, political and geographic environments. By defining the environment of the system, the dynamic context and all exchanges, influences and other factors are taken into consideration (Jackson *et al.*, 2010).

It is essential that every system has a purpose that is reflected in the identification of the function of the system (Jackson *et al.*, 2010). As a system is comprised of different elements, it is also important that each element also has a function. The system and its functions can have multiple functions. It can also be that multiple elements perform a single function together. When elements within a system are grouped together to perform a certain function, it constitutes a sub-system. The elements can be any combination of hardware, software, humans, processes and conceptual ideas. Sub-systems are positioned at a level appropriate to the function they perform (Jackson *et al.*, 2010).

The systems approach views the system of interest as interacting with other systems and consists of interacting sub-systems. It is important to consider all the interactions to design in all the inflows in, the intra-flows and the outflows from the system of interest (Jackson *et al.*, 2010).

When a system is synthesised, holistic methods are used to define the architecture of the entire system of interest. During the synthesis of the system, iteration is used to refine the system. When holistic methods are used, multiple system elements and their interrelationships are considered in the context of the whole (Jackson *et al.*, 2010).

The objective of a system is to solve a problem. To reach this objective, the systems approach considers the attributes of an entire system (Jackson *et al.*, 2010). In problem-solving systems, the main importance is not dealing with the entire system or where the boundary of the system of interest is drawn, but what proper entities and attributes of the system should be focused on (Chen, 1975). This means that in order to have an appropriate research product, the right problem needs to be identified. To do this, the key challenges underlying the problem need to be located by thinking through the problem and focusing on the critical elements.

2.1.2 The healthcare systems approach

To function as a system, every participating unit in the healthcare delivery system needs to recognise its dependence and influence on all other units (Reid and Grossman, 2005). To optimise the system as a whole, each unit must not only achieve high performance, but it is important that units must join together to optimise the performance of the system as a whole.

Reid and Grossman (2005) applied a systems approach to develop a four-level model of the patient-centred healthcare system. Reid and Grossman (2005) adapted the four levels of the healthcare system as described by Ferlie and Shortell (2001) to develop the model. The four nested levels of healthcare are the individual patient, the care team, the overall organisation and the environment in which the organisation is embedded.

Ferlie and Shortell (2001) realised the need for a multi-level approach for the delivery of quality care and for change. It is best to consider all the levels for the highest probability of successful change. This does not mean that every change effort should focus on all four levels, but rather that the levels that are focused on are considered within the context of the other levels. The multilevel approach to healthcare delivery change can be implemented top-down or bottom-up, incrementally or radically (Ferlie and Shortell, 2001). These levels are illustrated in Figure 2.1 and are briefly explained as followed:

1. **The patient level**

   Healthcare policies emphasise an increase of consumer-driven healthcare (Reid and Grossman, 2005). This means that healthcare delivery focuses more and more on the patient. This also means incorporating the values and wishes of patients into care processes (Reid and Grossman, 2005).

2. **The facility / care team level**

   The care team consists of individual physicians and groups of care providers whose collective efforts result in the delivery of care to patients (Reid and Grossman, 2005). The care team is the basic building block of a clinical microsystem. This is the smallest replicable unit within an organisation that contains within itself the necessary resources to do its work. Care is standardised where possible, but care is also customised to meet individual needs of patients (Reid and Grossman, 2005). For effective care,

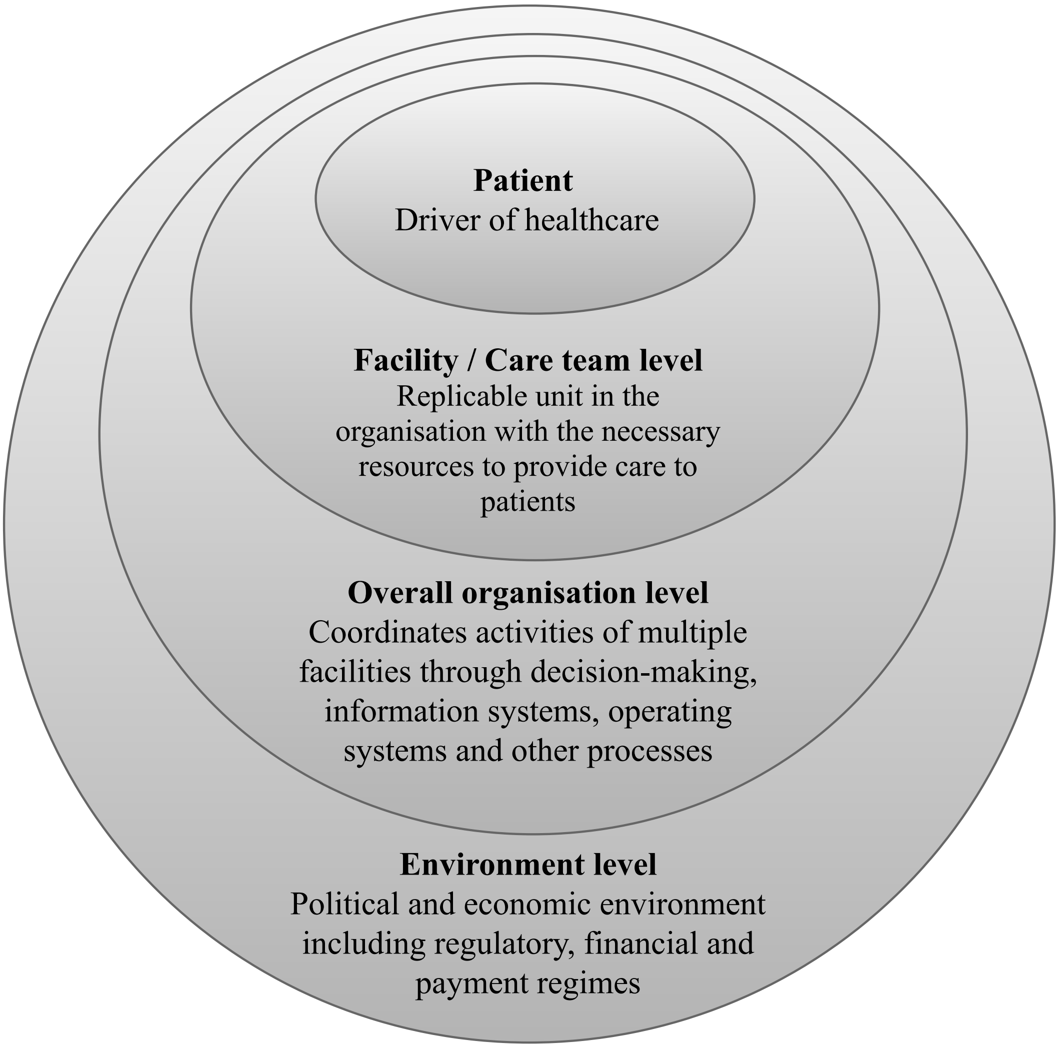


Figure 2.1: Healthcare system levels

physicians must have on-demand access to critical clinical information and administrative information.

3. **The overall organisation level**

   The organisation level provides infrastructure and other complementary resources to support the work of the care teams and microsystems (Reid and Grossman, 2005). The organisation coordinates the activities of multiple care teams and supporting units through decision making systems, information systems, operating systems and processes like financial, administrative, human-resources and clinical processes.

4. **The environment level**

The final level is the environment (Reid and Grossman, 2005). Reid and Grossman (2005) defines it as the political and economic environment that includes regulatory, financial and payment regimes. It also includes entities that influence the structure and performance of healthcare organisations directly. These entities also influence all other levels indirectly through their influence on the organisation.

2.2 Data management and healthcare

To understand data in the context of the healthcare sector, it was necessary to first gain an understanding of data management in general. In this section data management is defined (Section 2.2.1) and the importance of having a data strategy that is applicable to the relevant domain is described (Section 2.2.2). After a good understanding of data management in general was established, data management in the context of healthcare was also described (Section 2.2.3). This included reviewing traditional healthcare data management and the management of big data in healthcare.

2.2.1 Defining data management

Data is a valuable resource that needs to be managed through the process of creating, obtaining, transforming, sharing, protecting, documenting and preserving of data (O'Neal, 2012). Drucker (1988) stated that information is "data endowed with relevance and purpose." Before raw data is of any beneficial use, it must first be integrated with other data and transformed into information that guides decision-making (Dallemule and Davenport, 2017).

O'Neal (2012) further elaborates that data management comprises all the disciplines related to managing data. This includes the development and execution of architectures, policies, practices and procedures that manage the full data life cycle. The different aspects of data management are file naming conventions, policies and practices on how to create metadata and how to do documentation for the long term. The accuracy, completeness and security of data is ensured through data management. The different data management components include (O'Neal, 2012)):

- Data governance
- Data quality
- Master data management
- Metadata management
- Data architecture

- Privacy/security
- Data retention and archiving

2.2.2 Data strategy for beneficial data usage

It is also important to take the data strategy into consideration to understand what the data will be used for and how to enable that use (Dallemule and Davenport, 2017). A data strategy ensures the improved execution of how data is acquired, stored, managed, shared and used (Levy, 2018). The data strategy should be actionable, measurable, and relevant (Zaino, 2017). Another main objective of the data strategy is that it aligns and prioritises data and analytics activities with key organisational priorities, goals and objectives.

The data strategy strives to ensure that all data resources are positioned to be used, shared and moved easily and efficiently (Levy, 2018). Data is a critical asset that enables processing and decision-making. The data strategy ensures data is managed as an asset. To do this, the data strategy provides goals and objectives for the efficient and effective use of data. It also establishes common methods, practices and processes to manage, manipulate and share data in a repeatable way (Levy, 2018).

Normally, healthcare tends to have a defensive data strategy, because it operates in highly regulated environments where data quality and protection are of the utmost importance (Dallemule and Davenport, 2017). The defensive strategy focuses on minimising downside risks. That means that it includes activities such as ensuring compliance with regulations, detecting and limiting fraud and preventing theft (Dallemule and Davenport, 2017). Compliance with regulations includes rules governing data integrity and data privacy. Another objective that the defensive data strategy strives to achieve is to ensure the integrity of data flowing through the internal system of the company. This is done by identifying, standardising and governing authoritative data sources in a Single Source of Truth (SSoT) (Dallemule and Davenport, 2017).

Striking the balance between a defensive or offensive data strategy is important. The environment in which healthcare operates demands a defensive strategy, but activities of an offensive strategy are also needed (Dallemule and Davenport, 2017). Data needs to be standardised and uniform for a defensive strategy to comply with regulations and to implement data-access controls, but it also needs to be flexible to convert it into useful information through data analytics, modelling, visualisation, transformation and enrichment (Dallemule and Davenport, 2017). Therefore, one logical repository that contains one authoritative copy of all the important data is important (Dallemule and Davenport, 2017). This enables robust data provenance and governance controls which are required in healthcare. Furthermore, this controlled data should be flexible so that it can be managed to give it relevance and purpose. A successful data strategy includes all the different disciplines within data management

(Levy, 2018). The different elements of the two different data strategies can be seen in Table 2.1.

Table 2.1: Data strategies (Dallemule and Davenport, 2017)

	Defense	**Offense**
Key objectives	Ensure data security, privacy, integrity, quality, regulatory compliance, and governance	Improve competitive position and profitability
Core activities	Optimise data extraction, standardisation, storage, and access	Optimise data analytics, modelling, visualisation, transformation, and enrichment
Data management orientation	Control	Flexibility
Enabling architecture	SSoT	MVoT

2.2.3 Healthcare data management

After data management was explained in general, it was possible to also gain an understanding of data management in the context of healthcare. Therefore, this section briefly describes healthcare data management. Traditional healthcare data management is described first (Section 2.2.3.1), followed by a short description of big data in healthcare (Section 2.2.3.2) and lastly, healthcare data management on different system levels is described (Section 2.2.3.3).

2.2.3.1 Traditional healthcare data management

In order for a traditional healthcare data management system to function well, it needs an overarching system of governance, which enables the effective functioning of the rest of the data management components which includes: (i) data collection; (ii) data storage; (iii) ensuring data quality; (iv) data processing and analysis; (v) data dissemination; and (vi) the use of data (World Health Organization, 2008, 2014). Healthcare data management means that data related to the delivery of healthcare to patients is collected (Yang *et al.*, 2015), stored (Vreeland *et al.*, 2016), shared and used (Yang *et al.*, 2015).

Healthcare data management should be underpinned by clear legislative, regulatory and planning frameworks (World Health Organization, 2008, 2014). This legal framework should specify the roles and responsibilities of producers and users of healthcare data. Healthcare data management policies should be based on the principles of accountability and transparency, and should make explicit provision for the assurance of the ethical use of data and protection of individual privacy and confidentiality (World Health Organization, 2014). These frameworks manage the personnel, financing, logistics support, information and communications technology and coordination mechanisms with regard to healthcare data management (World Health Organization, 2008).

A wide range of policies and processes is needed to ensure data quality (World Health Organization, 2008). Data should meet high standards of reliability, transparency and completeness. Data is systematically and regularly assessed for quality by an independent data verification mechanism that is routinely conducted (World Health Organization, 2014). Other processes to improve the quality of data include regular local quality control and data-use checks, the use of clear definitions of data elements, up-to-date training and frequent feedback to data collectors and users (World Health Organization, 2008).

The collection of patient data such as diagnoses, treatments and outcomes is of key importance for good quality clinical care (World Health Organization, 2014). Within a traditional healthcare data management system, various healthcare data is collected, which includes patient data, such as: (i) patient demographics; (ii) encounter summaries; (iii) medical history; (iv) allergies; (v) intolerances; (vi) lab test histories (Ludwick and Doucette, 2009); (vii) individual patient records such as specific type of care data or data for conditions requiring long-term care and multiple visits; (viii) data collection for outpatient, admission and discharge registries (World Health Organization, 2014); and (ix) a core set of health indicators including health status, risk factors, service coverage and health systems indicators (World Health Organization, 2008, 2018).

The storage of patients' medical records is important for managing disease trajectory and clinical decision-making (World Health Organization, 2008). Healthcare data should be stored in an appropriate location where it is easily retrievable. Healthcare data storage should be well organised by: (i) restricting access to authorised users; (ii) coding the system to make records retrievable; (iii) following clear procedures for record distribution and refiling; and (iv) observing obligatory rules for the minimum period of maintenance and dispatch times (World Health Organization, 2008).

Data sharing means that healthcare data is shared and processed in a networked system (Vreeland *et al.*, 2016). There is a broad range of data users at different levels of the healthcare system (World Health Organization, 2008) and data is disseminated to facilitate the use of healthcare data (World Health Organization, 2014). Healthcare data should be shared securely. Secure data exchange means confidentiality, integrity, availability and timeliness of health and patient data. Data emanating from numerous care providers like family physicians, specialists, social workers, pharmacists, radiologists, dietitians, physiotherapists and nurses, is shared (Ludwick and Doucette, 2009). Stored patient data is shared electronically with authorised healthcare providers anytime, anywhere to support high quality care (Ludwick and Doucette, 2009).

As raw healthcare data has limited use, it needs to be processed and analysed to be of beneficial use (World Health Organization, 2008). The most important part of data processing and compilation is the extracting and integrating of data. Value is added to the source data through the process of extraction and transformation. This is done by: (i) removing mistakes and correcting for missing data; (ii) providing documented measures of degree of confidence in data; (iii) capturing the flow of transactional data for safekeeping; (iv) adjusting data from multiple sources to allow it to be used together; (v) structuring data to be usable by end-user tools; and (vi) tracking all these actions to tangibly support data quality assessments. To analyse data, it needs to be reviewed, processed, integrated with data from other sources and then have appropriate techniques applied to it (World Health Organization, 2014). Data produces meaningful insights after it is compiled, managed and analysed (World Health Organization, 2008).

Some usages of healthcare data include: (i) the standardisation of doctors' work practices; (ii) improving data availability; (iii) improving safety and quality of care (Vreeland *et al.*, 2016); and (iv) the incorporation of data into decision-making processes to enhance the operational efficiency between hospitals and for resource allocation (World Health Organization, 2014). Stored patient data is used to track patient medical history, interventions, encounters, lab test results as well as managing allergies and drug contraindications (Ludwick and Doucette, 2009).

2.2.3.2 Big data in healthcare

Big data refers to large, complex data sets that surpass the capabilities of traditional data management systems to store, manage and process data timely and economically (Nambiar *et al.*, 2013). Historically, the healthcare industry has generated large amounts of data through record-keeping, for compliance and regulatory requirements and for patient care (Raghupathi and Raghupathi, 2014). Big data in healthcare refers to electronic health data sets that are too large and complex to manage with traditional software and hardware, or by means of traditional or common tools and methods.

Big data in healthcare entails the management of very diverse and large volumes of data at a very high speed (Raghupathi and Raghupathi, 2014). Big data in the healthcare industry is made up of the totality of data related to patient healthcare. This includes clinical data, clinical decision support systems, patient data in electronic patient records, machine generated and sensor data and less patient-specific data like emergency data, news feeds and articles in medical journals (Raghupathi and Raghupathi, 2014).

Big data in healthcare enables the discovery of deep knowledge and value for the delivery of the best evidence-based, patient-centric healthcare (Yang *et al.*, 2015). Big data analytics in healthcare has the potential to improve care, save lives and lower cost through discovering associations and understanding patterns and trends within the data (Raghupathi and Raghupathi, 2014). Through the associations, patterns and trends that the synthesis and analysis of big data reveals, more thorough and insightful diagnoses and treatments can be developed. This results in higher quality care at lower costs and in better overall outcomes.

Big data analytics in healthcare has the potential to improve operational efficiencies, help predict and plan responses to disease epidemics, improve the quality of monitoring of clinical trials (Raghupathi and Raghupathi, 2014; Nambiar *et al.*, 2013), improve healthcare research and development (Raghupathi and Raghupathi, 2014), enhance evidence-based medicine (Yang *et al.*, 2015; Raghupathi and Raghupathi, 2014), incorporate genomic analytics, preadjudicate fraud analysis (Raghupathi and Raghupathi, 2014) and optimise the healthcare spending at all levels of the healthcare system including patients, hospital systems and governments (Nambiar *et al.*, 2013). Big data analytics also helps to make the shift from cure to preventive health.

A four-step process needs to be followed to utilise big data in healthcare (Bahri *et al.*, 2018). The process is depicted in Figure 2.2. Raw data is processed and analysed to assist decision-making. The first step is the generation of vast amounts of data from various sources. These sources include internal company data from its information system, Internet of Things (IoT) data, internet data and biomedical data (Bahri *et al.*, 2018).

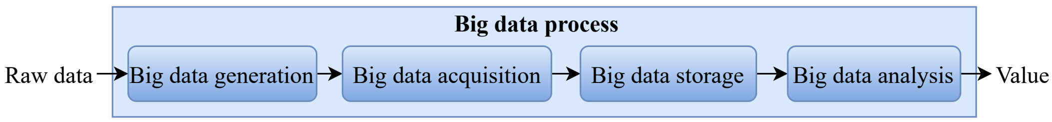


Figure 2.2: Big data process steps

The second step is data acquisition that can be subdivided into three substeps. These sub-steps are big data collection, big data transmission and big data preprocessing. Big data collection has to do with the acquisition and retrieval of vast amounts of raw data. These data can be structured, semistructured and unstructured and is retrieved from different sources like information systems, mobile devices, the IoT and open data. The information system is the centralised data warehouse that contains all the information about the activities of the organisation (Bahri *et al.*, 2018). The second sub-step is transmission which has to do with the transfer of data from the different data sources into storage management systems where it is processed and analysed (Bahri *et al.*, 2018). The last data collection step is big data preprocessing that ensures efficient and enhanced data for storage and analysis (Bahri *et al.*, 2018). To do this, redundant, noisy, incomplete and useless data is eliminated and the data is also integrated with other data for additional value. This also decreases the storage requirement and improves analytical accuracy (Bahri *et al.*, 2018).

To store big data, databases that are capable of handling vast amounts of data of various types and formats are used (Bahri *et al.*, 2018). These databases should be able to guarantee data security, availability and reliability. Data is stored for further analysis and processing.

Big data analysis is the most important process step (Bahri *et al.*, 2018). This is the step where value is generated as an output. Techniques and technologies are applied to mine and extract meaningful and valuable insights and hidden information from the large amount of processed and stored data (Bahri *et al.*, 2018).

Adopting Big data in public health does not come without ethical challenges and it is important to recognise the potential risk and unintended consequences (Vayena *et al.*, 2015). National and international legislation and guidelines were developed for very different historical conditions and are not effective or appropriate in addressing the new ethical challenges that Big data in healthcare poses. Through Big data much personal data can be accessed to be utilised for good, but the rights of the individual, such as the right to privacy, should still be respected. Big data can lead to misleading and inaccurate findings which may result in harm to individuals, businesses or communities (Vayena *et al.*, 2015). One such example is if a person is falsely identified as affected by an infectious disease. Harm can include financial loss, stigmatisation and the infringement of individual freedoms.

2.2.3.3 Healthcare data management on different system levels

According to Reid *et al.* (2005) information and information exchange are crucial to the delivery of care on all levels of the healthcare delivery system. Data is used at different levels of the healthcare system for healthcare service and system management (World Health Organization, 2008). The users of health data include those that deliver care to patients and those who are responsible for managing and planning health programmes. Therefore, the health data should be presented and disseminated in the appropriate formats to all audiences.

Care providers and care teams need access to at least three types of clinical information for diagnosis and treatment. These include the health record of the patient, the medical-evidence base and provider orders that guide the process of patient care (Reid *et al.*, 2005). The demand for timely sub-national data on service access, coverage, and quality for annual health reviews and operational planning processes are also increasing (World Health Organization, 2014).

At the organisational level, clinical, financial and administrative information is required for the measurement, assessment, control and improvement of the quality of their operations. The clinical, financial and administrative information needs to be captured, integrated and analysed to improve the quality and efficiency of healthcare. Interoperability and data-interchange standards enable the internal integration of the organisation's clinical, financial and administrative information. This also enables the organisation to link its system with other external care providers, insurers, vendors and regulatory bodies. Data domains that need to be standardised are administrative data, clinical data, medical images, prescription data and medical device data (Reid *et al.*, 2005).

The broader health data management system brings together data from multiple sources such as from health facilities, surveillance systems and other administrative data sources (World Health Organization, 2014). Facility-based data is an integral component of this broader system. This data is derived routinely and includes data collection from service delivery records and patient-provider interactions. They are also generated from a range of interventions offered such as treatments administered and treatment outcomes. This data can be used for various purposes, which includes managing patient care, epidemiological surveillance, monitoring of intervention specific programmes and quality assessments. Facility-based data can also be used for monitoring and evaluation on a country level and can be used to measure performance indicators of outputs, outcomes and impact.

As there is an increasing emphasis on monitoring and evaluation, there is also an increase of indicators, but this also increases the workload of health system staff and often does not lead to improvement (World Health Organization, 2014). Therefore, the number of indicators should be kept to a minimum and should consist of those that are used for decision-making on the various system levels. Countries should decide on key indicator sets for national planning, monitoring and evaluation of programmes, and for facility level management. At the facility level, indicators should be chosen to monitor efforts to improve health system performance.

The World Health Organization (WHO) has specified a list of 100 core health indicators, The Global Reference List, that serves as a normative guide for the selection of standard indicators that countries and partners' stakeholders can use for monitoring with respect to their priorities and capacity (World Health Organization, 2018). The Global Reference List reflects relevant indicators on universal health coverage, non-communicable diseases and other health-related environmental, social, economic and behavioural risk factors. The list comprises a selection of priority indicators across four domains, which include health status, risk factors, service coverage and health systems (World Health Organization, 2018).

Health status indicators consist of the core indicators which include mortality by age, sex and cause. Risk factor indicators relate to nutrition, environmental, behavioural, injuries and violence indicators. Service coverage indicators relate to the whole spectrum of health services which includes reproductive, maternal, newborn, child and adolescent, immunisation, HIV, TB, malaria, neglected tropical diseases, non-communicable diseases, mental health and substance abuse services. Lastly, health system indicators consist of indicators on health system inputs and outputs like health facility density and distribution, health workforce, health information, quality and safety of care, and health security capacity (World Health Organization, 2018).

2.3 The scope of the healthcare data management challenges

After laying the foundation of understanding the healthcare system and data management in the context of healthcare, it was important to understand the challenges that arise in the management of healthcare data in developing countries. A structured literature review was conducted to identify the healthcare data management challenges in developing countries. Many healthcare data management challenges exist in developing countries that impede the effective delivery of healthcare. The aim of this structured literature study is to gain a better understanding of the scope of challenges that currently exist and to determine which challenges are the most prominent ones. The results of this review were used to populate the challenge landscape for requirements specification in Chapter 3.

The text in Section 2.3 has been largely reproduced from a conference article with the following citation: van der Merwe, L., Bam, W. and De Kock, I., 2019, August. Healthcare Data Management in Developing Countries: A Systematic Review. In SAIIE NeXXXt.

This section begins with a description of the methodology followed to identify the healthcare data management challenges in developing countries (Section 2.3.1), which is followed by a quantitative analysis of the data management challenges (Section 2.3.2). Lastly, a discussion of the healthcare data management challenges in developing countries is presented (Section 2.3.3).

2.3.1 Structured literature review methodology to identify the healthcare data management challenges in developing countries

A structured literature review was conducted on 16 April 2019. The search followed the process described in this section. This study followed the systematic review process as described by Pickering and Byrne (2014). Firstly, the electronic search platform and the search terms that were used are described in Section 2.3.1.1. Secondly, in Section 2.3.1.2 the exclusion process is described from the initial search results to the eventual number of search results. The required data will be extracted as described in Section 2.3.1.3 which will be analysed to draw conclusions from. Some variations to the process as described by Pickering and Byrne (2014) are made in Section 2.3.1.3 as the steps in the process are inherent to how the study was carried out. The variation is explained in Section 2.3.1.3.

2.3.1.1 Search strategy

The first steps were to define the topic, formulate the research question, identify the keywords and identify and search databases (Pickering and Byrne, 2014). The topic was defined as "health care data management in developing countries". The research question was "What are the scope of health care data management challenges in developing countries?" The electronic search platform used for this structured literature review is Scopus. It was decided to use only Scopus, because it is the largest abstract and citation database of peer-reviewed literature that includes scientific journals, books and conference proceedings. Scopus consists of an extensive database with good quality and diverse sources which ensures good worldwide research coverage. Scopus developers claim that it is the largest single abstract and indexing database ever built (Burnham, 2006). With Scopus it is possible to enter multiple search requirements with different search operators. This allows Scopus to search for results in a very specific field, while maintaining objectivity with regard to the inclusion of all results in that field. Scopus was developed by Elsevier and the characteristics of both PubMed and Web of Science were combined. This allows enhanced utility for medical literature research and academic needs (Falagas *et al.*, 2008).

The search terms were chosen to cover the research field of healthcare data management challenges in developing countries, and to specifically include South Africa. To ensure the comprehensiveness of the initial Scopus search, different keywords were used for the same concept. The variants of the term "data management" for this study include "data processing", "data control" and "information management". These terms were used in inverted commas to return only articles where the words of the terms are used in conjunction. Furthermore, health care was searched as one word and as two words enclosed in inverted commas to accommodate articles with the spelling difference. The search terms "developing countr" and "south africa" were used to narrow the search down to only developing countries and South Africa. These search terms were used with an asterisk to include articles where variants of these search terms are used. Lastly, synonyms of challenges were used with asterisks to ensure comprehensive coverage of articles that identified healthcare data management challenges. These synonyms include problems and issues. Therefore, the relevant literature was retrieved with the following search keywords:

(TITLE-ABS-KEY ("Data processing" OR "data management" OR "data administration" OR "data handling" OR "data control" OR "information management") AND TITLE-ABS-KEY (healthcare OR "Health care")

AND TITLE-ABS-KEY (“south africa*” OR “developing countr*”) AND TITLE-ABS-KEY (challenge* OR problem* OR issue*))

2.3.1.2 Exclusion criteria

The next step was to assess the publications to ascertain if they were relevant and whether they should be included (Pickering and Byrne, 2014). The first exclusion criteria was to exclude articles that were published prior to 2008. Data management has evolved tremendously over the years and including articles that are too old will result in including irrelevant challenges in this study. This study strives to address relevant healthcare data management challenges of developing countries.

The second exclusion criterion was by document type. Books were excluded from this review.

Lastly, the abstracts of the remaining articles were read to determine whether they are truly relevant to this study field. Articles were excluded during this phase based on whether they clearly mentioned any healthcare data management challenges in the abstract. Sometimes, articles were included in the initial search because the study used data management to address a totally different challenge than what this study addresses. Such articles, for example, were excluded during this phase.

This method ensured that only publications that are truly relevant to healthcare data management challenges in developing countries were included. Figure 2.3 illustrates the process of starting at the initial search protocol, to how the search results were narrowed down to the eventual number of search results through the different exclusion steps.

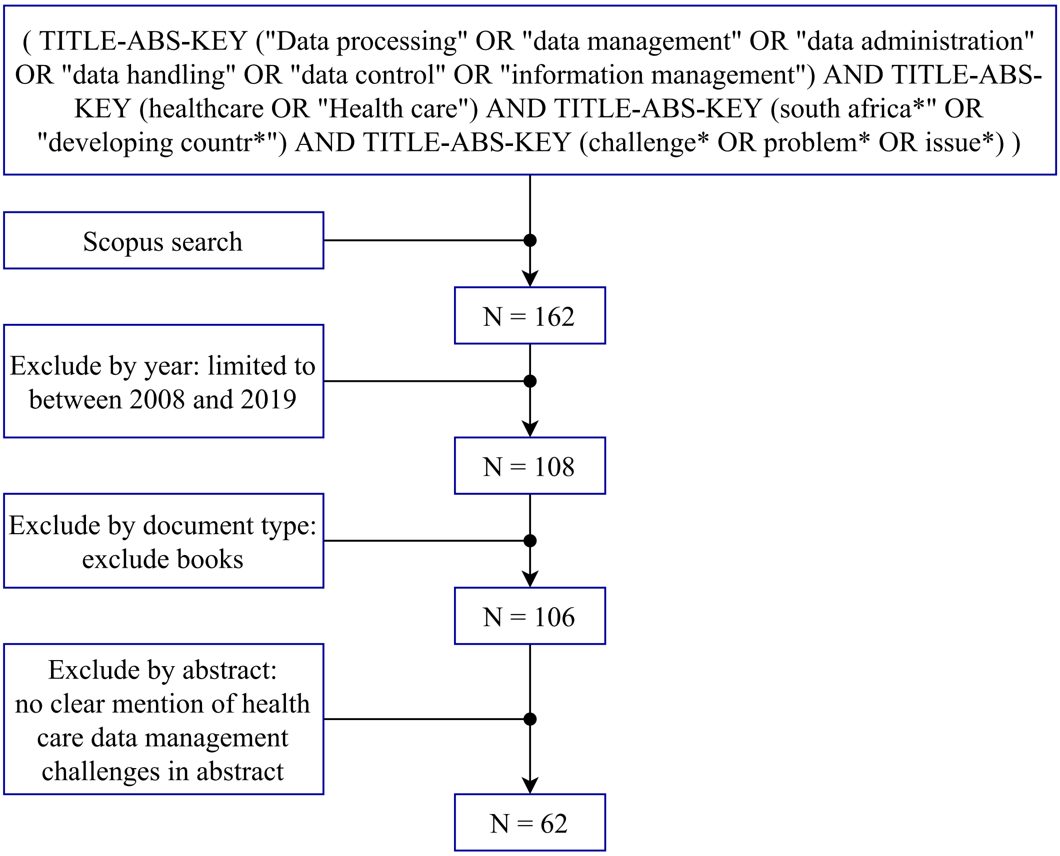


Figure 2.3: Structured literature review process

2.3.1.3 Data extraction

After it was determined which articles to include in the study, the next step was to develop the structure for a personal database on the topic (Pickering and Byrne, 2014). After the completion of the article selection process, data was extracted using MS Excel. Excel was used to develop a healthcare data management scope of challenges. This was done by creating an Excel spreadsheet. The aim of the scope of challenges was to identify all the possible healthcare data management challenges found in developing countries and to categorise them into appropriate categories and subcategories. This scope of challenges enabled the quantitative analysis of the extracted data. Other relevant data extracted from literature into the Excel sheet include:

- Year of publication
- Author(s)
- Article title
- Country the study focused on

Whenever an article mentioned a healthcare data management challenge of developing countries, an 'x' was marked under the category to which that challenge belongs. Articles addressed multiple challenges and some challenges fell under multiple categories. The categories changed and more categories were included as the data was extracted. This caused the scope of challenges to expand. At the end of the data extraction process, the categories were assessed to merge very similar challenges in the different categories. Challenges were classified under different data management components.

Due to the evolving nature of the scope of challenges, steps seven to 10 as described by Pickering and Byrne (2014) were automatically carried out in the reading of all the publications. As new challenges were found the structure of the database was updated until the final database was complete.

After the database was completed, it gave a holistic view of all the healthcare data management challenges that developing countries face. These challenges were categorised according to the different data management components. These data management components can be regarded as the main categories. Challenges were designated to subcategories under these categories. Other main categories are additional to data management components, but also have an impact on data management. The main categories that the challenges were extracted to in the Excel database are:

- Poor governance
- Integration challenges
- Data collection challenges
- Data storage challenges
- Data processing challenges
- Data transmission challenges
- Data retrieval challenges
- Data utilisation challenges
- Data monitoring challenges
- Data reporting challenges
- Data analysis challenges
- Data quality challenges
- Data security challenges
- Infrastructure and technology challenges

- Cost and financial support challenges
- Human factor challenges
- System implementation challenges

The subcategories under the categories were determined as they surfaced in the literature. When a new subcategory appeared in literature, it was added under the appropriate category. A subcategory, general challenges, was included under most categories if a challenge was mentioned without a specific regard of a subcategory.

2.3.2 Quantitative analysis of the healthcare data management challenges

Quantifying the number of times that challenges fell into each category gives an indication of which healthcare data management challenges are the most prominent. The total number of occurrences recorded for all the categories combined is 1 124.

For each category the challenges were further categorised into subcategories. This section will discuss the distribution of the challenges between the different categories and subcategories. Figure 2.4 illustrates the number of challenges that occurred in each category and Figure 2.5 illustrates the Top 10 subcategories across the different categories.

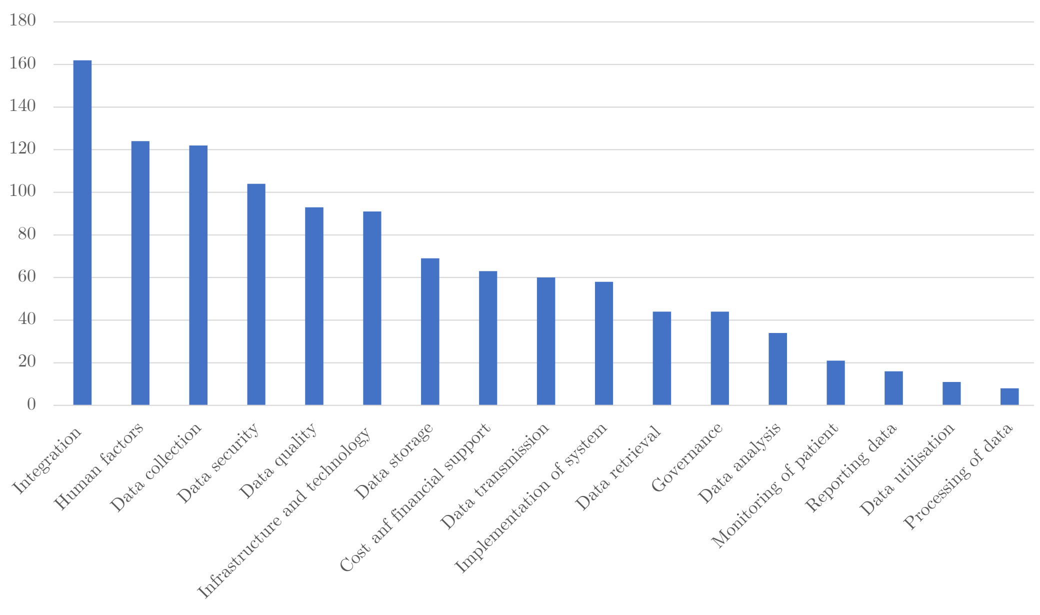


Figure 2.4: Number of occurrences per category

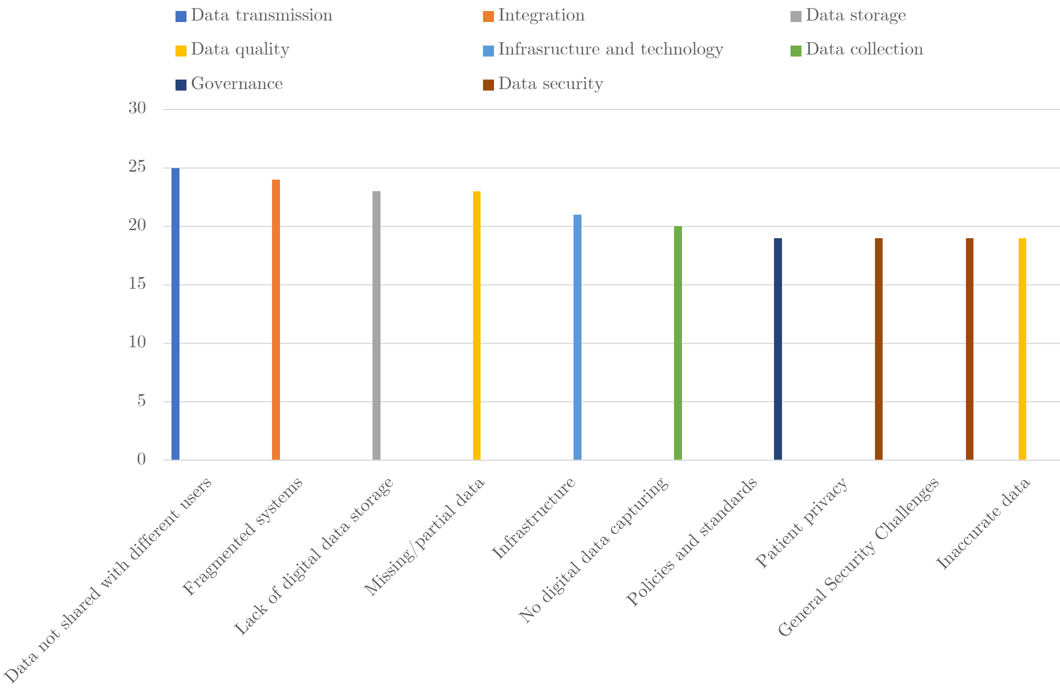


Figure 2.5: Top 10 most occurring subcategory challenges

As can be seen in Figure 2.4 the six categories with the most challenges are integration, human factors, data collection, data security, data quality, and infrastructure and technology challenges. Integration challenges are the highest by far, scoring 162 occurrences. Human factor challenges at 124 had the second most occurrences. The category with the third most challenges was data collection with 122 occurrences, followed by data security with 104. Data quality had 93 occurrences and technology and infrastructure had 91 occurrences.

It is clear from Figure 2.5 that data not shared between different users is the subcategory with the most occurrences. This is followed by fragmented systems. Lack of digital data storage and missing or partial data is tied in third place. Lack of infrastructure is the next subcategory with the most occurrences. From there on it is no digital data capturing, policies and standards, patient privacy challenges, inaccurate data and general security challenges.

It is interesting to note that the category data transmission did not have a high number of occurrences, but one of its subcategories, data not shared with different users, is the subcategory with the highest score. Out of the Top 10 data management subcategory challenges, categories data quality and data security featured twice, while the other categories featured only once in the Top 10. It is also interesting to note that integration challenges, the category with the most occurrences, also featured only once on the Top 10 subcategory list. All these findings convey the intricacy of the healthcare data management challenges in developing countries.

2.3.3 Discussion of the healthcare data management challenges

This section describes the different challenge categories as found in literature. The discussion is presented according to the ranking of healthcare data management challenges. The categories are presented in the order of the highest-scoring category to the lowest-scoring category from Section 2.3.3.1 to Section 2.3.3.12. Each category is introduced, followed by a discussion from literature of the different subcategory challenges of that category.

2.3.3.1 Integration challenges

In Section 2.3.2 it was found that integration challenges was the category with the most occurrences. Challenges that contribute to the complication of integrating healthcare data management systems were designated to this category.

There are many factors contributing to integration challenges. There are many stand-alone information systems (Masana and Muriithi, 2017) and different systems have heterogeneous forms that make them even more difficult to integrate (Bhaskaran *et al.*, 2013; Khan and Hoque, 2016*a*, 2017). Often the necessary standardisation and interoperability between systems are missing (Fritz *et al.*, 2015; Idoga *et al.*, 2018). Standardisation and interoperability will enable different systems to share data among them.

Data is complex (Sampath *et al.*, 2017), and to manage data effectively it needs to be organised and aggregated. Aggregated data can help to make better decisions based on holistic views of data (Aanestad *et al.*, 2014; Bhaskaran *et al.*, 2013).

Paper-based data management systems make integration basically impossible (Palmer and Simms-cendan, 2012; Turan and Palvia, 2014). Incompatible technologies further complicate integration (Wang *et al.*, 2010; Turan and Palvia, 2014).

Another challenge is that even when it is possible to integrate systems and share data, new security challenges arise (Idoga *et al.*, 2018). Integration or sharing data should not compromise the security of the data.

As stated in Section 2.3.2 it was found that fragmented systems is the second biggest challenge on the Top 10. It is also the integration challenge with the most occurrences compared to other subcategories. It was found that in many developing countries fragmented systems were due to different sub-systems operating separately and often even vertically (Mgozi and Weeks, 2011). Integration between governmental bodies or within the Ministry of Health is also missing (Turan and Palvia, 2014). Fragmented and heterogeneous systems trap data needed for proper decision-making (Khan and Hoque,

2017) and Sharifi *et al.* (2013) stated that fragmented and inaccessible clinical information have a negative effect on healthcare quality and cost.

2.3.3.2 Human factors

According to the analysis of Section 2.3.2 human factors is the second biggest challenge category. It is interesting to note that none of its subcategories are in the Top 10, but it is still the third biggest healthcare data management challenge category. All human-related data management challenges were designated to this category. The five most prominent human-related challenges are the low skill level of staff, the lack of training, the lack of staff, digital illiteracy and lack of participation.

Many healthcare workers do not have the required skills to do the data management tasks they need to do. They make mistakes with data collection, do not know how to use the routine health information system, and lack data security, data management and data analysis skills (Koivu *et al.*, 2016).

There is also a lack of training (Sharifi *et al.*, 2013). A lack of training means that healthcare workers are not trained before they start work and it also means that their skills will not improve in the future, because there are no training opportunities. Even in cases where there was training, there is no ongoing data management training available (Mate *et al.*, 2009).

Digital literacy is another significant challenge (Idoga *et al.*, 2018; Koivu *et al.*, 2016; Mgozi and Weeks, 2011). For healthcare workers to manage data, they need to be able to use the necessary information and communication technologies. Healthcare workers do not have the expertise to operate equipment such as tablets, computers, and smart phones (Warkulwiz *et al.*, 2014).

Not only is the skill level or the digital literacy of the staff low, there is also a significant shortage of staff (Haskew *et al.*, 2015). This includes a shortage of data capturers (Allorto and Wise, 2015). The shortage of staff influences data quality (Kaposhi *et al.*, 2014).

Another factor is that humans do not participate in the data management system as they should. Braa and Sahay (2012) found that sometimes health managers lack faith in the health management information system. A lack of collaboration of doctors and health personnel can lead to poor data quality (Turan and Palvia, 2014).

2.3.3.3 Data collection challenges

Data collection challenges had the third most challenges documented. The data collection challenges subcategories with the most occurrences are no digital data capturing, errors with collection, inefficient collection processes and methods, lack of proper entry forms, time constraint, duplication and different healthcare facilities collecting different data. It is also important to note that no digital data capturing was sixth on the Top 10 subcategory list.

No digital data capturing mostly entails that data is captured on paper and documentation is also done on paper (Paul *et al.*, 2009; Bindle and Joost, 2010). Sometimes there are systems that convert the paper-based data into digital format, but usually there is a lag before that is done (Monda *et al.*, 2012). Paper-based capturing makes it very difficult to exchange and access information and to monitor patients' progress (Masana and Muriithi, 2017). Masana and Muriithi (2017) also stated that the lag before data is digitalised prevents real-time data accessibility.

Often times handwritten data is illegible, affecting the quality of the data (Allorto and Wise, 2015). Data capturing errors also affect the quality of data. These errors include partial data collection or entry mistakes by the capturer (Mate *et al.*, 2009; Khan and Hoque, 2016*a*,*b*).

There is also a concern about the suitability of data entry forms (Sharifi *et al.*, 2013). Sometimes forms do not allow the entry of relevant data in the way that it is structured (Medhanyie *et al.*, 2017). There exists a need to standardise these forms to ensure all relevant data is captured (Allorto and Wise, 2015).

There are different data entry points. This causes the same data to be captured multiple times. This data is aggregated only later. Data is therefore duplicated by different data capturers (Braa and Sahay, 2012; Medhi *et al.*, 2012).

It has also been found that different healthcare facilities collect various types of data (Alkraiji *et al.*, 2016). Different sub-districts and many healthcare facilities have different interpretations of how data should be captured and managed (Kaposhi *et al.*, 2014).

Another data collection challenge is the limited time available to collect data (Nicol *et al.*, 2013; Turan and Palvia, 2014; Alkraiji *et al.*, 2016). In developing countries, it takes long to collect data and often data capturers do not have enough time to collect it. Overworked healthcare personnel have multiple responsibilities and do not have the time to collect data (Nicol *et al.*, 2013), especially when paper-based data capturing is used which is very inefficient and not time-efficient (Nyumbeka and Wesson, 2014).

2.3.3.4 Data security challenges

Data security had the fourth most challenges identified. Any challenge regarding the protection of data or protection against data breaches is included in this category. The subcategory with the most occurrences is patient privacy. General data management challenges also has just as many occurrences and both these subcategories share the seventh place on the Top 10 subcategory challenges list. Authorisation, confidentiality and data integrity are the other significant data security challenges.

Medical systems are more connected and networked. It is necessary to have identifiable health data in health data repositories across these systems for data accessibility and sharing, but the increase in connectivity and identifiable health data causes an increase in patient privacy risks and security breaches (Khan and Hoque, 2016*a*, 2017). Electronic documents have many privacy and security risks because they are accessed and transferred easily (Turan and Palvia, 2014). Unmanaged servers endanger data, bandwidth and other devices on the network (Whalen *et al.*, 2014).

Preservation of anonymity and security of patient records are significant concerns, but in some developing countries there are no privacy laws that protect identifiable data (Warkulwiz *et al.*, 2014; Idoga *et al.*, 2018). A need also exists for well-coordinated regulatory frameworks for proper governance of the privacy and security of patient health information (Mgozi and Weeks, 2011).

Another significant concern is unauthorised access (Idoga *et al.*, 2018). Authorisation services include policy management, role management and role based access control (Ganiga *et al.*, 2018), but sometimes even authorised users can use data maliciously (Bhaskaran *et al.*, 2013).

2.3.3.5 Data quality

Data quality is the fifth biggest data management challenge category and it has two subcategories in the Top 10 subcategory challenges list. Missing data is tied in third place and inaccurate data shares the seventh place. These are the two main data quality subcategories, but poor data quality in general was also often mentioned (Turan and Palvia, 2014; Fritz *et al.*, 2015; Jobson *et al.*, 2018). Other data quality challenges were data duplication, unstructured data and data discrepancy or inconsistency.

Quality data is important for many healthcare functions. Missing and inaccurate data impedes functions such as record-linking, proper diagnosis and analysis, patient monitoring and clinical and public health decision-making (Vlădescu *et al.*, 2009).

Causes of missing or inaccurate data are nurses that simply do not write out all patient data (Sharifi *et al.*, 2013), doctors that do not write out the diagnosis, but only symptoms and prescriptions (Aanestad *et al.*, 2014) and errors made by less qualified staff (Khan and Hoque, 2017).

2.3.3.6 Technology and infrastructure challenges

The technology and infrastructure challenges category is the sixth biggest challenge category. It involves all data management challenges related to the technologies and infrastructures needed for data management in healthcare. The highest-scoring subcategories were infrastructure challenges, network availability, software challenges and power supply challenges.

The infrastructure required for data management is lacking (Fritz *et al.*, 2015; Bijlmakers *et al.*, 2017; Medhanyie *et al.*, 2017; Alkraiji *et al.*, 2016).

Some of these infrastructures needed for data management include information technology infrastructure (Ganiga *et al.*, 2018), communication infrastructure (Wang *et al.*, 2010), and technological infrastructure to allow appropriate information storage and sharing (Idoga *et al.*, 2018). New technologies cannot be introduced to healthcare data management systems because the current infrastructure is unable to support them (Aanestad *et al.*, 2014).

Network availability is also a big challenge. Ganiga *et al.* (2018) found that there is a great need for high speed internet connectivity, but often the internet connection is unreliable. The unreliability of the internet forces lower level sub-systems to be completely paper-based (Bergum *et al.*, 2017). Internet signal strength is often weak (Nyumbeka and Wesson, 2014) and bandwidth costs are high, inhibiting the health systems' effectiveness (Mgozi and Weeks, 2011).

Power supply has been found to be erratic (Idoga *et al.*, 2018). In many resource-limited areas healthcare facilities do not have a reliable electricity supply to support data management activities such as electronic data storage (Aanestad *et al.*, 2014; Whalen *et al.*, 2014).

2.3.3.7 Data storage challenges

Data storage has the seventh most challenges documented. Challenges are designated to this category if it posed challenges to the storage of data. The highest data storage subcategory, lack of digital data storage, is ranked second on the Top 10 list for subcategory challenges. Loss of data is the second biggest data storage concern. Other data storage subcategories include unstructured storage, inefficient storage, reliability of storage and the lack of storage infrastructure.

Many healthcare systems in developing countries are still managing data in traditional paper-based systems (Medhi *et al.*, 2012; Ganiga *et al.*, 2018; Vivanco and Oropeza, 2017). This makes it difficult to access patient information (Patra *et al.*, 2012) and gives rise to challenges in data aggregation, transmission and analysis (Medhi *et al.*, 2012). This paper-based stored healthcare data is also located at different geographical locations. This limits the adoption of a system-wide approach to healthcare management (Turan and Palvia, 2014). For some manual paper-based systems health statistics that are recorded in log books are sent to regional offices for data capturing of metrics into a centralised database, but these log books are sent infrequently (Cline and Luiz, 2013).

Loss of data is also a common data storage challenge. Data that is needed for patient care or programme management gets lost (Letebo and Shiferaw, 2016). In paper-based data storage the loss of health record books is common (Irawan *et al.*, 2016) and in some paper-based systems the record room only retains records of patients for five years. This results in the loss of continuous patient data (Ganiga *et al.*, 2018). Chimbari (2017) found that data sets are collected at high cost, but are not analysed and get lost over time. Digital data storage is prone to loss of data in the case of hardware or software failure. Efficient backup is needed to prevent loss of data (Whalen *et al.*, 2014).

2.3.3.8 Cost and financial support challenges

The cost and financial support challenges category represents the eighth biggest healthcare data management challenge. All the different cost aspects of healthcare data management are categorised under this category. Implementation cost was mentioned the most in literature. Other data management cost aspects mentioned include infrastructure costs, technology costs, systems costs, data storage costs and training costs.

The cost of implementing healthcare data management systems is the biggest cost challenge and is the main reason why developing countries struggle to adopt digital healthcare data storage (Letebo and Shiferaw, 2016). Cost is also the reason why existing digital medical records are not integrated into information technology (Bhaskaran *et al.*, 2013). Additional costs are incurred for acquiring, installing and maintaining equipment for integration (Idoga *et al.*, 2018).

The expensiveness of technology is also a great concern. Healthcare companies have developed technologies to improve healthcare data management, but the problem is they are too expensive for developing countries (Mukhopadhyay *et al.*, 2018). Healthcare data management software and hardware are expensive (Sharifi *et al.*, 2013), as is equipment to monitor patients at healthcare facilities (Nyumbeka and Wesson, 2014).

2.3.3.9 Implementation of systems challenges

Implementation of systems is ranked ninth from the different challenge categories. Challenges regarding the implementation of systems is designated to this category. Its main subcategories are resistance to change, training, technical infrastructure available to support implementation and the existing culture. The cost of implementing systems is the greatest implementation challenge, but that is categorised under the category cost and financial support.

Healthcare system users sometimes have negative perceptions about the usefulness and the associated threats regarding new systems. This causes a resistance to adopt the new systems (Turan and Palvia, 2014). Threats regarding privacy and security make organisations reluctant to adopt new systems (Mgozi and Weeks, 2011). One example of resistance to change due to perceived usefulness is that reports from radiology are handwritten. Staff objected to typing these in electronically, because it takes too long and workloads are too high (Aanestad *et al.*, 2014). Patients also resist change. Warkulwiz *et al.* (2014) found that patients did not want their data to be entered electronically, and preferred paper-based systems instead.

The availability of technical infrastructure is very important for systems implementation (Fritz *et al.*, 2015), but poor infrastructure makes this very difficult (Whalen *et al.*, 2014). To ensure adoption of widespread implementation, currently available infrastructure should be used to support the implementation of systems, rather than having to develop new and sophisticated technologies to support implementation (Wang *et al.*, 2010).

Another factor is that the need for training makes implementation difficult (Allorto and Wise, 2015). When new systems are implemented, it involves training which is very expensive (Sharifi *et al.*, 2013). After implementation, ongoing training of personnel is also needed (Mate *et al.*, 2009).

2.3.3.10 Governance challenges

According to Section 2.3.2, governance challenges is ranked tenth. All challenges that relate to the governance of healthcare data management is categorised in this category. Its main subcategories are policies, legislation and standards. Other subcategories include regulations, frameworks and leadership.

Policies are needed at a national level to ensure coherence (Idoga *et al.*, 2018), but there is a lack of adequate policies and procedures (Alkraiji *et al.*, 2016). Some common policies that are missing are policies regarding security (Turan and Palvia, 2014), Information and Communications Technology (ICT) (Mengiste, 2013) and well-defined access to policies to ensure authentic users can access data (Patra *et al.*, 2012). Policies for the use of information systems are also important, for they grant access to the systems to authorised users anywhere and at any time (Palmer and Simms-cendan, 2012). Kaposhi *et al.* (2014) recommended changes in knowledge translation, data verification, programme management and standardisation policies.

Having the necessary standards is also very important to have a system-wide approach to patient healthcare management (Turan and Palvia, 2014). It is found that many developing countries do not have informational and care standards and they have limited existing regulations (Aanestad *et al.*, 2014). Digital health record standards have also been adopted only recently (Gainer *et al.*, 2012).

It was also found that there is an absence of legislation regarding data management (Idoga *et al.*, 2018). Health data is collected from public and private hospitals, but because the required legislation is not in place, this data cannot be shared or used (Turan and Palvia, 2014). Turan and Palvia (2014) also found that there is no legislation safeguarding personal health information.

2.3.3.11 Data transmission challenges

Data transmission was quite a low-scoring challenge category, but one of its subcategories, data sharing with different users, had the highest number of occurrences of all the subcategories. Challenges regarding the transmission of data from one place to another by whatever means were included in this category. Other data transmission challenges are difficulties with disseminating data to patients, data transfer latency, unreliable network availability for transmission and transmission errors. In some cases, data was transferred in paper format.

A key challenge in developing countries is to make healthcare data accessible from rural to urban (Ganiga *et al.*, 2018). Paper-based systems only allow data to be accessed from one place (Mvelase *et al.*, 2015). Therefore, a need exists to digitalise data for data accessibility and data sharing, but if digital systems are not linked or integrated, data accessibility and exchange to different users will still pose a significant challenge (Masana and Muriithi, 2017).

Some clinical information technologies do not allow data-sharing between clinicians, labs, hospitals, pharmacies and patients (Pahl *et al.*, 2015) and data is not shared between different levels of healthcare either (Haskew *et al.*, 2015). Different healthcare facilities use their systems with their own localised data networks. This inhibits data-sharing and the adoption of a system-wide approach to patient healthcare management (Turan and Palvia, 2014). Data is isolated in silos which impedes data-sharing between care providers (Bhaskaran *et al.*, 2013). Legislative frameworks and an unwillingness of companies to share data are yet some other factors contributing to the lack of data-sharing (Mgozi and Weeks, 2011).

2.3.3.12 Data retrieval challenges

Data retrieval ranked twelfth. Its subcategories consist of challenges regarding the retrieval and accessibility of data. The main subcategories are historical data retrieval, real-time data retrieval, the lack of remote data accessibility and timely data accessibility.

Paper-based and manually filed medical records are difficult to access and impede good service to patients (Kaseke *et al.*, 2017). Electronic databases can make retrieval of stored data easier (Letebo and Shiferaw, 2016). Data retrieval is important to make accurate clinical decisions (Bhaskaran *et al.*, 2013) and inaccessible clinical information has a negative effect on quality of healthcare (Sharifi *et al.*, 2013).

Many developing countries use systems that do not allow timely data accessibility. For instance, healthcare systems that are divided into different categories which operate on different systems makes access to patient data very difficult (Idoga *et al.*, 2018). Other examples are processes that are error-prone and time-consuming cause a delay in data accessibility (Mvelase *et al.*, 2015).

Recently, there has been an increase in the need to access data from remote locations (Paul *et al.*, 2009). This is needed so that doctors can view patient health records from anywhere and give advice and treatment (Patra *et al.*, 2012).

2.4 Conclusion on healthcare data management

This chapter gave a description of healthcare and data management in healthcare. It described the healthcare delivery as a system with different system levels. This chapter gave insight on how the healthcare system functions with its components such as its elements, functions, levels and environment. The healthcare system has four nested levels which are the individual patient, the care team, the overall organisation and the environment. The environment of the system also has an influence on the system. This gave a good idea of the structure of the healthcare system.

Data management was explained in general and it was further discussed in the context of healthcare. The different data management components were identified and the importance of a data strategy was discussed. The data strategy ensures the improved execution of how data is acquired, stored, managed, shared and used. A defensive data strategy was fitting for the healthcare context. This means downside risk should be minimised through compliance regulations, detecting and limiting fraud and preventing theft. Healthcare data management was also described with regard to traditional data management and the management of big data in healthcare. The different functions of data management in the context of healthcare were identified. The discussion on traditional and big data management also gave insights into the change that is occurring in healthcare data management and the transition from traditional to new data management systems that is taking place.

Furthermore, the scope of healthcare data management challenges in developing countries was investigated to determine the whole range of challenges and to determine which challenges are the most persistent. The six significant healthcare data management challenge categories identified in this chapter were integration challenges, human factors, data collection challenges, data security challenges, technology and infrastructure challenges, and data quality challenges.

This gives an idea of the scope and complexity of the healthcare data management challenges in developing countries. To help solve healthcare data management challenges in developing countries, the challenges identified in this study should be taken into consideration. Therefore, in the next chapter the scope of challenges is used to derive a challenges landscape for healthcare data management in developing countries. This enabled the specification of requirements that the possible research product should adhere to, to address the healthcare data management challenges satisfactorily.

Chapter 3

The requirements specification of healthcare data management in developing countries

In the previous chapter the healthcare data management scope of challenges was identified. There are multiple challenge categories and subcategories to take into account. The scope of challenges gave an indication of all the different aspects to be taken into account when dealing with the lingering healthcare data management problem in developing countries. To address all these challenges as a whole, the different requirements that a possible research product must adhere to must be specified. To gain a better understanding of the needed design requirements, the data value chain was derived from literature. In doing this, data management across its whole value chain is understood and the significant components of data management are identified. It also gives a more comprehensive view of data management as a system with different components. Constructing the data value chain also conveys the importance of the beneficial use of data and what value adding components are needed to enable the beneficial use of data. Then, based on the value chain and the scope of challenges, the challenges landscape of healthcare data management was developed which was used to specify the requirements that research products should adhere to, to help improve the healthcare data management problem in developing countries. The objectives of this chapter are therefore to understand the healthcare data value chain with which to develop the challenge landscape that enables the definition of the design requirements (Section 3.1) and to define the design requirements to address the healthcare data management problem in developing countries (Section 3.2). Lastly, this chapter on requirements specification is concluded (Section 3.3).

3.1 Developing the healthcare data value chain and challenges landscape

This section describes how the healthcare data value chain and challenges landscape were developed. A framework with different system components was needed to enable the categorisation of the different healthcare data management challenges according to the different components of the healthcare system. The categorisation of the challenges allows for an in-depth understanding of the challenges according to their corresponding healthcare data management system components, which enables the requirements specification for this study. The framework of Franz *et al.* (2015) is appropriate for this as it identifies generic system components applicable to any domain which allows the categorisation of the healthcare data management challenges.

The framework of Franz *et al.* (2015) was developed for the energy industry and describes the components of the energy market system, but the components included in this framework are applicable to any system, irrespective of the application domain, as all systems consist of these components. Therefore, the components described by Franz *et al.* (2015) can be applied to the domain of healthcare data management too. The purpose of the framework of Franz *et al.* (2015) is to analyse energy market systems and develop effective interventions for the challenges faced in the different energy market systems. The framework of Franz *et al.* (2015) consists of a generic model that describes the components of the energy market system across three levels and derives interventions for these components on the different levels. As these levels and components are generic, they can be applied in the context of healthcare data management. Similar to the framework of Franz *et al.* (2015), healthcare data management consists of a value chain with multiple levels that function as a system. Both the framework of Franz *et al.* (2015) and this study strive to seek supporting interventions for their specified challenges. Therefore, the framework of Franz *et al.* (2015) is appropriate to determine the healthcare data management value chain for this study, as well as determine the challenges landscape across the whole healthcare data management value chain.

The framework of Franz *et al.* (2015) was used as a guide to identify the different components needed in the healthcare data value chain. The framework of Franz *et al.* (2015) consists of two stages. The first stage is mapping the energy market system and stage two is the design of supporting interventions. Stage one was used to assist the development of the healthcare data value chain that was developed by investigating current big data value chains. The similarities and differences of the different proposed value chains were used to complement each other in the development of the healthcare data value chain that is used for this study (Section 3.1.1). Stage two of this framework was used to develop the challenges landscape on the value chain (Section 3.1.2).

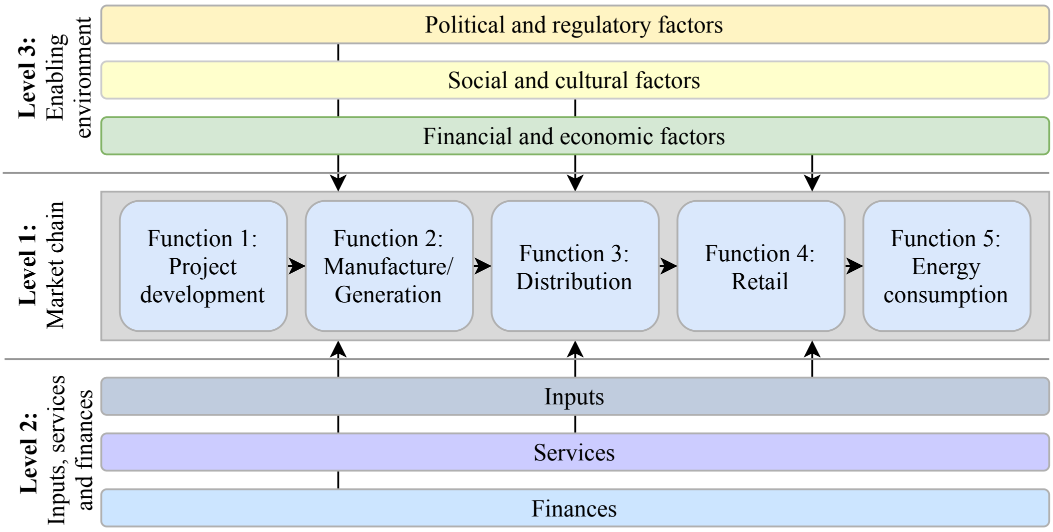


Figure 3.1: The different levels of the generic value chain (Franz *et al.*, 2015)

3.1.1 Healthcare data value chain

Using the framework by Franz *et al.* (2015), the healthcare data value chain was mapped. The energy market value chain framework describes value chains of energy systems, but the principles of the components can be used to represent the healthcare data value chain with its different levels. Energy market functions can be related to the data value chain components.

Franz *et al.* (2015) developed the energy market system analytical framework to categorise each energy market and a set of processes to determine how each energy market operates and what obstacles are prevalent in that energy market. The energy market system is the entire system delivering energy services, including all its component parts, subdivided into three levels. The generic energy market chain system is illustrated in 3.1.

Level one is the energy market or value chain. The market chain is at the centre of the market system. It describes how the products move through the chain along which it is converted, from the primary generators to the final end user (Franz *et al.*, 2015). The different stages add value to the products. The market chain has been divided into broad functions. The broad functions are the most important stages of delivery of each energy product or service (Franz *et al.*, 2015).

At level two all the secondary value chains that support the main value chain are included (Franz *et al.*, 2015). Level two value chains comprise inputs, services, and finances. To be able to carry out the market chain functions effectively and efficiently, access to a variety of specific secondary inputs, services and finances are needed (Franz *et al.*, 2015).

Level three is comprised of the enabling environment which includes the

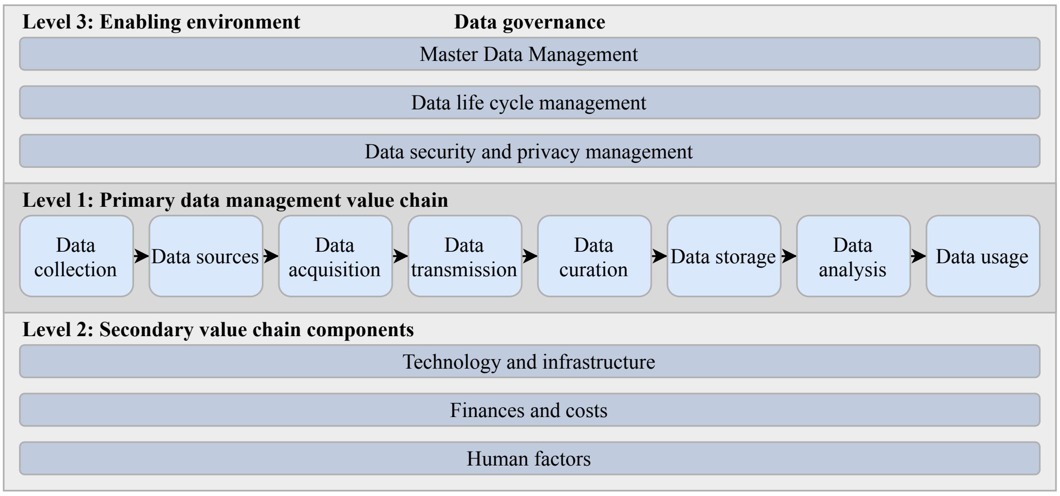


Figure 3.2: Healthcare data value chain

diverse overarching enabling environment factors that shapes how the market chain and level two components operate (Franz *et al.*, 2015). The enabling environment gives the conditions under which the market chain operates. The factors that are included in level three are political and regulatory, social and cultural, and financial and economic factors (Franz *et al.*, 2015). Such an environment comprises policies, regulations and cultural practices. It is often generated by institutions such as national and local authorities and research agencies (Franz *et al.*, 2015).

The healthcare data value chain was developed based on the energy market chain and by using different big data value chain representations. The similarities and differences of the big data value chain representations were used complementary to create a healthcare data value chain for the purpose of this study. Big data value chains were used as described by Labrinidis and Jagadish (2012), Miller and Mork (2013), Gandomi and Haider (2015), Cavanillas *et al.* (2016), Bhadani and Jothimani (2016) and Wang *et al.* (2018). Big data value chains were used, because no documented value chains could be found in literature on "traditional" data value chains. Literature on big data value chains was reviewed to develop the value chain as it describes "traditional" data value chains to an extent. These big data value chains focused on adding value to the data for the purpose of analysing and using it beneficially. The value chain was created to plot the challenge landscape, which in turn aided the specification of the design requirements of the research product. The value chain can be seen in Figure 3.2.

The primary value chain consists of all the functions that add value to the data from where it is collected as raw data to where it can be used beneficially. These functions have to do with the transformation of the data. The secondary value chain is all the inputs and services needed to realise the primary value chain including technology and infrastructure, finances and cost and human contributions. The enabling environment for this value chain consists of the different components of data governance as the enabling environment. The different healthcare data management value chain levels are described in the following sections. The primary data value chain is described first (Section 3.1.1.1), followed by the description of the secondary value chain (Section 3.1.1.2) and lastly, the enabling environment is discussed (Section 3.1.1.3).

3.1.1.1 Primary data value chain

This section describes the primary data value chain. It was derived from the combination of different big data value chains. Literature on big data value chains was used to develop the data value chain for this study. The primary data value chain starts with the data collection function where raw data is gathered from the real world in structured, semi-structured and unstructured formats. This raw data is gathered into various sources depending on the type of data. Data from these sources is then acquired for centralised storage where it can be analysed. The acquired data is transmitted to storage, but needs to be curated first to ensure the quality of data. If analysis of data is performed on bad quality data it will result in bad quality analysis. The curated data is then stored centrally where analysis can be performed to gain valuable insights. The analysis of the data is used to assist in making decisions, support in operations and perform appropriate actions.

1. **Data collection**

   Data is divided into three formats (Wang *et al.*, 2018). They are structured data such as traditional electronic healthcare records (EHR), semi-structured data such as the logs of health monitoring devices and unstructured data such as clinical images (Wang *et al.*, 2018). Data that can be organised using predefined data models is structured data (Bhadani and Jothimani, 2016). Structured data includes tabular data in relational databases. Unstructured data cannot be organised by these predefined models and includes video, text and audio (Bhadani and Jothimani, 2016). Semi-structured data fall between the categories of structured and unstructured data. These clinical data needs to be collected from different internal or external locations. Data is collected from the real world and depending on the content format, it is then stored in the appropriate databases (Wang *et al.*, 2018). These databases are the data sources where data is collected to. For example, patient information must be collected into a healthcare information system and forms part of structured data. With this collection of data, human collectors are sometimes used to collect data digitally from paper-based forms. Semi-structured data like logs of health monitoring devices are automatically generated to the data source and do not have to be collected by people.

2. **Data sources**

   Data sources include all sources necessary to provide insights for daily operations and to solve business challenges (Wang *et al.*, 2018). Financial data, administrative data, patient behavioural data, population data, medical device data, clinical information system data, electronic health record data, radiology information system data, laboratory information system data (Cavanillas *et al.*, 2016) and data from patient records (Wang *et al.*, 2018) are data that is found in structured data sources. Other data sources include mobile apps, social media, home care sensors (Wang *et al.*, 2018), picture archiving and communication systems (Cavanillas *et al.*, 2016), and electronic medical images (Bhadani and Jothimani, 2016). These are the sources where the data is acquired from for analysis (Bhadani and Jothimani, 2016).

3. **Data acquisition**

   The primary goal of data acquisition is to read data provided from various communication channels, frequencies, sizes, and formats (Wang *et al.*, 2018). During this step the data is obtained from all the different data sources (Bhadani and Jothimani, 2016). The aim is to gather data from distributed information sources so that it can be filtered, and cleaned before it is stored for analysis (Cavanillas *et al.*, 2016).

4. **Data transmission**

   Distributed data that is acquired is transferred to a data storage and processing infrastructure where it is processed and analysed. Transfer is carried out in two phases. During the first phase data is transferred from the source to the data centre, and from there it is transferred within the centre (Bhadani and Jothimani, 2016). The data centre not only stores data, but also helps with the acquisition, organisation and management of data.

   Data transmission also means that different entities share healthcare data in a networked system (Vreeland *et al.*, 2016). Authorised care providers like family physicians, specialists, social workers, pharmacists, radiologists, dietitians, physiotherapists, and nurses share healthcare data to support high quality care (Ludwick and Doucette, 2009).

5. **Data curation**

   Data quality requirements need to be met for its effective usage (Cavanillas *et al.*, 2016), but the data acquired from various data sources may be redundant, noisy and inconsistent (Bhadani and Jothimani, 2016). Therefore, the data is processed to improve its quality. Good quality data improves the accuracy of the analysis and reduces storage expenses.

Data curation steps include integration, cleaning and elimination of redundant data (Bhadani and Jothimani, 2016). This phase is responsible for ensuring that data is trustworthy, discoverable, accessible, reusable, and fits its purpose (Cavanillas *et al.*, 2016).

During this step data is moved, cleaned, split, translated, merged, sorted, and validated (Wang *et al.*, 2018). This means that structured data that is typically contained in an eclectic medical record might be extracted from healthcare information systems and converted into a specific standard data format, sorted by the specified criterion, and then the record is validated against data quality rules before it is loaded into a storage system for further processing and analysis (Wang *et al.*, 2018).

6. **Data storage for analysis**

   Data storage systems should provide reliable storage space and powerful access to the data (Bhadani and Jothimani, 2016). The storage system should consider the factors of consistency, availability and partition tolerance. Database transactions should be guaranteed, but the storage system should be flexible when data volumes and complexity grow (Cavanillas *et al.*, 2016). The data storage principles are based on compliance regulations, data governance policies and access controls (Wang *et al.*, 2018).

7. **Data analysis**

   Data analysis is the step where data is changed to make it usable in decision-making (Gandomi and Haider, 2015; Cavanillas *et al.*, 2016). During data analysis, data is explored, transformed, and modelled with the goal of highlighting relevant data. Data analysis adds structure to data to support decision-making and domain-specific usage (Cavanillas *et al.*, 2016). Useful hidden information is synthesised and extracted this way. Cluster analysis and regression analysis are two traditional data analysis methods (Bhadani and Jothimani, 2016). During data analysis the provenance between the input and results should be maintained, as well as the metadata, so that other analysts can recreate the analysis (Miller and Mork, 2013).

8. **Data usage**

   Visualised results obtained from the data analysis are used to determine the necessary action. The analysed data is used to support decision-making (Miller and Mork, 2013). Outputs such as various visualisation reports, real-time information monitoring, and meaningful business insights obtained from data analysis are generated for decision-making (Wang *et al.*, 2018). Reporting allows data to be visualised in useful ways to support daily operations and help to promote faster, better decision-making (Wang *et al.*, 2018).

3.1.1.2 The secondary value chain

The literature on the big data value chains did not address the secondary value chain components as described by Franz *et al.* (2015). As it was important to obtain a comprehensive view of the challenges across the data management system, the secondary value chain components that were added include: (i) technology and infrastructure; (ii) finances and cost; and (iii) human contributions. They were derived as components to the data management system as they surfaced as challenge categories in the scope of healthcare data management challenges, but are not part of the primary value chain and also do not fit under level three of the energy market system as described by Franz *et al.* (2015). They were components that were added to the healthcare data value chain in order to obtain the comprehensive challenges landscape that was constructed in Section 3.1.2.

1. **Technology and infrastructure**

   These are all the technologies and infrastructure needed as an input to support the primary value chain from the collection of data through to the usage of the data. There are many technologies and infrastructures needed, but their details are not important to this study and therefore, they are only mentioned on a high level. It is still important to address them as a component of the data management system as they were addressed as a challenge category in the scope of challenges by various sources of literature (Fritz *et al.*, 2015; Bijlmakers *et al.*, 2017; Medhanyie *et al.*, 2017; Alkraiji *et al.*, 2016; Ganiga *et al.*, 2018; Nyumbeka and Wesson, 2014; Idoga *et al.*, 2018).

2. **Finances and costs**

   As with technology and infrastructure, this component was added as it falls under the description of the secondary value chain described by Franz *et al.* (2015) and it relates to the challenge category finances and cost that were identified in the scope of challenges (Sharifi *et al.*, 2013; Ganiga *et al.*, 2018; Yaqoob *et al.*, 2017; Haskew *et al.*, 2015; Fritz *et al.*, 2015). It is the component that includes all of the financial inputs to the primary value chain from data collection to data usage.

3. **Human contributions**

   The last level two component identified is human contributions. Human interactions also play a significant role in the value-adding process of the value chain from the collection of data, all the way through to its usage. It was included in this study for the same reason as the previous two components. It falls under level two of energy market systems as described by Franz *et al.* (2015) and human factors was a challenge category identified in the scope of challenges (Idoga *et al.*, 2018; Koivu *et al.*, 2016; Mgozi and Weeks, 2011; Braa and Sahay, 2012; Allorto and Wise, 2015). Therefore, human contributions was included in the data management system as a level 2 component. It includes all the contributions humans make to the value chain from collection to usage of data.

3.1.1.3 The enabling environment

Data governance is the enabling environment for the healthcare data management system as described by the literature on big data value chains (Wang *et al.*, 2018; Cavanillas *et al.*, 2016). Data governance describes how to harness data in the organisation (Wang *et al.*, 2018).

1. **Master data management**

   Master data management includes the processes, governance, policies, standards, and tools for managing data and ensuring that it is properly standardised, removed, and incorporated. This creates the immediacy, completeness, accuracy, and availability of master data that supports data analysis and decision-making (Wang *et al.*, 2018). Sound healthcare data management depends on organised processes for gathering, sharing, analysing and using health data (World Health Organization, 2008, 2014). This is achieved when countries, institutions and management structures adopt and adapt health data standards.

2. **Data life-cycle management**

   Data life-cycle management is the process whereby business information is managed throughout its life cycle. The different phases include archiving data, maintaining the data warehouse, testing and delivering different application systems and deleting and disposing of data (Wang *et al.*, 2018).

3. **Data security and privacy management**

   Data security and privacy management is the platform that provides enterprise-level data activities. This includes discovery, configuration assessment, monitoring, auditing, and protection. Organisations face ethical, legal, and regulatory challenges due to the complexity involved in data management. Data rules and control mechanisms are needed for highly sensitive data. This prevents security breaches and protects privacy. Suitable policies, standards and compliance requirements that restrict the permission of users ensure a safe environment for the proper use of information (Wang *et al.*, 2018).

   An example of an international standard that provides requirements for establishing, implementing, maintaining and continually improving data security management is ISO/IEC 27001. It strives to preserve the confidentiality, integrity and availability of data.

> An example of a comprehensive data protection act in the South African context is the Protection of Personal Information Act (POPIA). It aims to give effect to the constitutional right to privacy, but at the same time balancing this against competing rights, such as the right to access information and interests. Similarly, the General Data Protection Regulation (GDPR) is a legal framework that sets guidelines for the collection and processing of personal information from individuals in the European Union.

3.1.2 Healthcare data management challenges landscape

The challenges landscape was developed by using the developed healthcare data value chain and the scope of challenges identified in Chapter 2. The value chain was developed according to stage one of Franz *et al.* (2015). Stage two is the identification and analysis of potential supporting interventions (Franz *et al.*, 2015). During this stage all the main obstacles are identified within the market system and within the different levels. The challenges identified in the scope of challenges can be related to the obstacles described by Franz *et al.* (2015) that can be allocated to the different levels and components of the healthcare data management system. This forms the challenge landscape that contributes to the enablement of the specification of requirements to address the healthcare data management problem in developing countries.

All the challenges identified in the scope of challenges were mapped on the value chain and placed under its different levels and components. This placement of the challenges in this way helps to identify the requirement of addressing the challenge at a specific place in the value chain. The challenge landscape can be seen in Appendix B. In the next section the developed challenge landscape is used along with prior knowledge discussed in previous sections to specify the requirements of the framework.

3.2 Requirements to address the healthcare data management problem in developing countries

In order to develop a framework that contributes to addressing the challenges of healthcare data management, the requirements that it should adhere to must be specified. Requirements were extracted from the chapters on healthcare data management, its scope of challenges for developing countries and the challenge landscape. The requirements specification ensures that the proposed research product responds to the stated healthcare data management problem appropriately, and that the study aim is reached. Van Aken and Berends (2007) describe requirements under five categories:

- Functional requirements (FR): the performance demands on the object to be designed. One such requirement is that the realisation of the research product should solve the problem.

- User requirements (UR): the specific requirements from the viewpoint of the user. Ideally, they explain the constraints and how the research product should be used. Examples of such requirements are that people presently working in the system should be competent to work in the new system and to use the new tools and procedures. One more example is that the system should be user-friendly.

- Design restrictions (DR): entails the preferred solution space. This places restrictions on the proposed research product and specifies its limits and exclusions. For instance, it should not take longer than a certain amount of time, it should not require more than a certain amount of money and the realisation of the research product should entail as few changes as possible to the current system.

- Boundary conditions (BC): requirements that should be addressed unconditionally. These requirements may include that the system should comply to legal requirements, ethical habits, code of conduct, business policies and fit with the current business culture.

- Attention points (AP): the requirements that are relevant to the development of the research product, that should be noted as desirable and should be considered, but do not have to be met.

These categories were used to categorise the requirements. Different requirements were extracted from Chapters 2 and 3. The various requirements are listed in Section 3.2.1 to Section 3.2.5. In each table the different requirements were identified according to a specific requirements category.

3.2.1 The functional requirements specification

Specifying the functional requirement ensures that the core performance demands that the framework should meet, are considered. Addressing the specified functional requirements ensures that the problem is addressed comprehensively. Table 3.1 specifies the functional requirements that state the core performance demands that ensure the problem is met comprehensively.

Table 3.1: Functional requirements specification

ID	Requirement	Motivation	Source section
FR1	The framework should enable the improvement of healthcare data management that improves healthcare delivery to patients and improves care delivery management and decision-making	Data is a valuable resource with relevance and purpose that is applied to aid the delivery of care to patients and to help the management and operations of healthcare activities	2.2.1, 2.2.2, 2.2.3
FR2	The framework should be able to be used to give an as-is assessment of healthcare data management	In order to be able to introduce improvements to healthcare data management, the as-is state should firstly be determined to determine what the appropriate improvements are from that state. The scope of challenges and challenge landscape identified all the different challenges under different components that can be included under the as-is state from where improvements should begin	2.3, 3.1.2
FR3	The framework should describe the incremental improvement of healthcare data management	There are many challenges that can arise in an organisation's healthcare data management system which cannot be addressed individually. Healthcare data management improvements should be done incrementally and across the whole scope of the data management system to ensure all parts are improved	2.3, 3.1.2
FR4	The framework should consider all the important data management components relevant to healthcare	There are many important components of data management that the framework needs to include to ensure that healthcare data management is improved comprehensively and to address the different challenges identified in the scope of challenges and the challenges landscape	2.2.1, 2.3.3, 3.1.1

Continued on next page

Continued from previous page

ID	Requirement	Motivation	Source section
FR5	The framework should include necessary system components to describe healthcare data management as a functioning system	Healthcare data management should be seen as a system with interacting elements that work together to realise a goal. The combination of the interacting components should be improved to ensure improvement. Improvement in only one element of the system will not ensure that the whole system is improved. To ensure improvement in healthcare data management, it must be addressed as a system that must be improved in its totality	2.1.1, 2.1.2, 2.2.3.3, 3.1.1
FR6	The proposed research product should include the necessary technological and infrastructural components needed for healthcare data management	Many of the challenges that occur in healthcare data management are due to the lack of the necessary technological and infrastructural components to carry out effective healthcare data management. These components need to be included in the proposed research product to ensure their absence does not prevent the improvement of healthcare data management	2.3.3.6, 3.1.1.2, 3.1.2

3.2.2 The user requirements specification

The user requirements were determined from the viewpoint of the user. The requirements of Table 3.2 were stated to ensure the framework is developed for the intended user and that it is usable by that user.

Table 3.2: User requirements specification

ID	Requirement	Motivation	Source section
UR1	The framework should be usable to managers or change agents of healthcare entities to enable data management improvement on a strategic level	The healthcare data management has a very broad range of challenges that indicates that the problem needs to be addressed on a strategic level. To address the overall challenges of healthcare data management, the overall aims and the means of achieving them should be included in the framework development. The framework should be developed for users with a managerial and strategic mindset	2.3, 3.1.2
UR2	The framework should be generic so that it is usable to different national healthcare entities	The scope of challenges investigated healthcare data management challenges of different developing countries. Because challenges from different countries were identified, the proposed research product should be generic enough to be applicable to different countries and help address their challenges, but specific enough to give valuable insights	2.3
UR3	The framework should be user-friendly and intuitive	The scope of challenges identified ease of implementation and resistance to change as reasons why new improvements fail. The users should find it straightforward to use the framework so that they do not get discouraged about using it	2.3

Continued on next page

Continued from previous page

ID	Requirement	Motivation	Source section
UR4	The framework should use standard domain language to be easily understandable	Domains have a standard language that describe the components of that domain. The proposed research product should be written in such a way that users can easily understand it without having to worry about ambiguity. The meaning of words should be clear. Data management and healthcare terminology should be clear and their standard language should be used	2.2

3.2.3 The boundary conditions specification

Boundary conditions should be met unconditionally. The requirements specified in Table 3.3 ensure that the framework addresses all the necessary legal and ethical considerations that the framework should comply with.

Table 3.3: Boundary conditions specification

ID	Requirement	Motivation	Source section
BC1	The proposed research product should consider governmental and national policies, acts and regulations	The regulatory environment has an influence on how the data of the organisation can be managed. The regulatory environment should be considered in the design of the model	2.2.1, 3.1.1
BC2	The proposed research product should describe the ethical considerations around healthcare data management	There are many ethical considerations to take into account when dealing with health data. The surge of big data also introduces many new ethical challenges. The ethical challenges should be recognised and measures should be taken to address them in the proposed research product	2.2.3.2

Continued on next page

Continued from previous page

ID	Requirement	Motivation	Source section
BC3	The proposed research product should incorporate the privacy and security of data	To manage healthcare data, a defensive strategy is usually followed. One of the key objectives of a defensive strategy when managing data is to ensure data privacy and security. Data security and privacy is exceptionally important in healthcare to prevent data theft and malicious use of data and should be strictly addressed	2.2.2, Table 2.1, 2.3.3.4, 3.1.2
BC4	The framework should include the necessary internal standards and policies	Policies and standards are essential to the governance of data management. They specify how the data management activities should be carried out, are principles to guide data management activities and are implemented as procedures and protocols. Data management activities are very ineffective without the necessary policies and standards to guide their execution	2.2.2, 3.1.2

3.2.4 The design restrictions specification

The design restrictions determine the preferred solution space. Restrictions that are placed on the framework were specified and the limits and exclusions placed on the framework are documented in Table 3.4.

Table 3.4: Design restrictions specification

ID	Requirement	Motivation	Source section
DR1	The framework is limited to data management in the healthcare domain	The scope of challenges and value chain are specific for healthcare data management and are not applicable to other domains. The problem space that the framework addresses is in the healthcare data management domain only	2.2, 2.3, 3.1.2
DR2	The proposed research product should not specify specific technologies, processes and methods to achieve its goal. The framework does not elaborate on the technical details of healthcare data management systems	The improvement of healthcare data management should not be restricted to certain technologies and infrastructure. Specifying certain technologies and infrastructure can limit the advancement of healthcare data management. Being bound to certain technologies and infrastructure can inhibit the ability to improve	2.3.3.6
DR3	The framework should describe different stages of healthcare data management improvement without prescribing how to achieve the improvements	Prescribing a set improvement roadmap can lead to excluding the best improvement methods, especially regarding best practices and technology that are constantly improving and can lead to the proposed research product being outdated very soon. The proposed research product should rather describe the improved state and it should be left to the user to determine how the improvements will be reached	2.3.3.9

3.2.5 The attention points specification

The attention points are relevant requirements that are desirable and should be noted and considered, but do not have to be met as a mandate. These attention points were specified in Table 3.5.

Table 3.5: Attention points specification

ID	Requirement	Motivation	Source section
AP1	The desired focus of the framework is to assist the improvement of healthcare data management in developing countries, but does not have to be limited to developing countries	The scope of challenges focuses on the healthcare data management challenges in developing countries and does not comprehensively describe challenges of developed countries. Developed countries can benefit from the model if they struggle with similar challenges and if their healthcare data management system is similar to that of a developing country	2.3
AP2	The proposed research product should strive towards the standardisation of all data management components across different units	It should be the focus to standardise the whole healthcare data management system, but it is very unlikely that this will be achieved fully. It should be desired that all data management systems should be integrated, but as this might not be completely possible, the proposed research product should include additional methods to achieve this goal	2.2.2
AP3	It should be considered that best practices are evolving	Technology is constantly evolving. The proposed research product should guide the user to consider the current best practices when improvements are sought	2.2.3.2

3.3 Conclusion on the requirements specification

The healthcare data value chain was developed in this chapter to describe healthcare data management comprehensively across the whole value chain as a system with interacting components and to enable the identification of the challenges landscape. The scope of challenges was used to populate the healthcare data value chain which is the base of the challenge landscape. The challenges landscape contributed to the enablement of the specification of the different requirements needed to address the healthcare data management problem in developing countries across the whole value chain.

The requirements specified in this section enable the development of a research product that can help address the persistent healthcare data management problem through meeting these requirements. These requirements must be taken into consideration when trying to find a research product that will help improve the healthcare data management problem appropriately.

In the next chapter maturity models are described as a research product to help address the healthcare data management challenges. The maturity model is described to illustrate their use and how it can be applied to the context of healthcare data management. In Chapter 5 a maturity model is developed that meets the specified requirements of this chapter that ensures that the developed model helps address the specified problem.

Chapter 4

Maturity models

In the previous chapter an exposition of the healthcare data management challenges landscape was given to illustrate the persistent challenges in healthcare data management across the whole value chain. This was used to determine the relevant requirements that the proposed research product should adhere to in order to ensure that the proposed research product satisfactorily deals with the stated problem. In this chapter maturity models (MM) are proposed as a suitable research product to help address the healthcare data management challenges. Firstly, the origin, purpose and value of maturity models are explained for a better insight into their origin and what their usability is (Section 4.1). Secondly, the basic structure of maturity models is described to gain a good understanding of what components are needed to make a maturity model useful (Section 4.2). Thirdly, the importance of using a defined methodology, design decisions and design principles is stated, along with examples of well-established design methodologies (Section 4.3). Literature on maturity models in the context of healthcare data management were studied to determine how maturity models have been applied in the past in this context, what can be learnt from previous studies and how this study can be an improvement on existing maturity models (Section 4.4). This chapter on maturity models is concluded in Section 4.5.

4.1 The origin, purpose and value of maturity models

In this section the origin of maturity models is described (Section 4.1.1), as well as their purpose of use (Section 4.1.2). This background information on maturity models explains why maturity models originated, what maturity models are used for, and how their use has evolved over time to be of more beneficial use.

4.1.1 The origin of maturity models

The use of reusable models that encapsulate concepts that are common to many enterprises is very popular due to the complexity of business and information systems engineering (Mettler, 2010*b*). These models are called reference models. A reference model constitutes a reference for a certain domain through including reusable state-of-the-art practices (Rosemann, 2003; vom Brocke, 2007). Reference models that focus specifically on the evolution of systems are known as maturity models (Mettler, 2010*b*). Maturity models are based upon the assumption of predictable patterns of organisational evolution and change (Röglinger *et al.*, 2012). Maturity models typically represent theories about how the capabilities of an organisation evolve in a stage-by-stage manner along an anticipated, desired or logical path (Röglinger *et al.*, 2012). From the 1970s multiple different maturity models started to be developed and following the introduction of the Capability Maturity Model (CMM) in the early 1990s, the popularity of maturity models increased considerably. It is concluded that maturity models have a long history of improving systems progressively.

4.1.2 The purpose and value of maturity models

Maturity models are designed to assess the maturity of a selected domain based on a set of criteria (De Bruin *et al.*, 2005). Maturity models give guidance to the gradual development process of systems by incorporating formality into the improvement activities (Mettler, 2010*b*). They are artefacts that determine the *status qua* of the capabilities of an organisation and derive measures to improve from there (Becker *et al.*, 2009). In other words, the basic purpose of maturity models is to outline the stages of maturation paths. This includes the characteristics of each stage and the logical relationship between them (Röglinger *et al.*, 2012).

It is important to assess the positioning of an organisation regarding its capabilities for continual improvement (Becker *et al.*, 2009). A maturity model is a supportive tool that assesses the as-is state of an organisation, derives and prioritises improvement measures and thus controls the progress of their implementation (Becker *et al.*, 2009). The maturity model provides criteria and characteristics that are necessary to be fulfilled to reach a particular maturity level and which serve as the scale to assess the position of the organisation on the evolutionary path regarding the specified criteria (Becker *et al.*, 2009). Therefore, the usefulness of maturity models is evident.

4.2 The basic structure of maturity models

This section deals with the basic constructs of maturity models that have been established over time. This was done to determine what the basic constructs of maturity models are, to ensure all essential components of maturity models are well understood and that all components are applied appropriately when the model was developed.

Fraser *et al.* (2002) stated that all maturity models share a number of common components. Maturity models define a number of dimensions (domain components) at several stages of maturity, with a description of characteristic performance at the different levels of granularity. Basic elements of maturity models include a number of maturity levels, a number of dimensions (domain components) and a number of elements or activities for each dimension (domain component). Each maturity level has a descriptor and a generic description. The generic description of each level is a summary of the characteristics of each level as a whole (Mettler, 2010*b*). Each element or activity also has a description of itself as it might be performed at each level of maturity (Mettler, 2010*b*). Some studies use the term dimensions for the concept of domain components. However, the term "dimensions" can refer to the dimensions of many different constructs and can be misleading. Therefore, the term "domain components" is used in this study to avoid confusion.

The rest of this section describes the different types of maturity models (Section 4.2.1), describes what is meant by the domain, domain components and capability areas of maturity models (Section 4.2.2) and explains what maturity levels are (Section 4.2.3).

4.2.1 Maturity model types

There are three distinct maturity model types that can be distinguished depending on the structure and complexity of the model (Fraser *et al.*, 2002). Maturity grids are the simplest form of maturity model that consists of text descriptions for each activity at each level of maturity. CMM-like models are the most sophisticated type of maturity models and are based upon a particular architecture more formal than that of maturity grids. General descriptions of maturity for each level of evolution are precisely described, but individual descriptions for each activity at each level are not stated. These models are often accompanied by extensive supporting materials. The last type is Likert-like questionnaires which are hybrids in between these two extremes (Fraser *et al.*, 2002).

4.2.2 Domain, domain components and capability areas

Maturity models can focus on a defined domain or they can have a general focus and they do not have a specific domain that they focus on (De Bruin *et al.*, 2005). When a maturity model focuses on a specific domain, it targets that domain and is applied to it. A focused domain determines the specificity and extensibility of the model.

A stage-gate approach enables the differentiated maturity assessment within complex domains and is achieved by additional layers of detail (De Bruin *et al.*, 2005). The additional layers enable the separate maturity assessment for a number of individually separate and distinct areas. The layers are represented by the domain, domain components and sub-components. Domain components are significant and independent particular parts of a specific domain that play a significant role in the maturity of the domain. Domain sub-components are distinct capability areas within the domain components. These distinct capability areas allow further detail that enables targeted maturity level assessment and improvement (De Bruin *et al.*, 2005). Domain sub-components and capability areas are used interchangeably, but the term capability areas will be used from here onwards.

The layered model within a domain enables a deeper understanding of relative strengths and weaknesses within the domain and therefore, more specific improvement strategies can be applied (De Bruin *et al.*, 2005). Domain components and capability areas are critical for complex domains and they should be mutually exclusive and collectively exhaustive (De Bruin *et al.*, 2005).

Maturity within a domain can also be represented by a series of one-dimensional linear stages which provide an 'average' maturity stage for the entity under study (De Bruin *et al.*, 2005). It gives a simple means of comparing maturity stages, but does not sufficiently represent maturity within complex domains. Hence, the layered model with a domain, domain components and capability areas is used for a more detailed maturity assessment (De Bruin *et al.*, 2005).

Although domain components and capability areas add more detail and give a deeper understanding of the strengths and weaknesses in distinct capability areas, it is important to keep the number of domain components and capability areas low (De Bruin *et al.*, 2005). The reason for this is to minimise perceived complexity and to ensure the independence of the components.

De Bruin *et al.* (2005) recommends an extensive literature review to identify domain components in a mature domain. In a mature domain the availability of great amounts of literature and tested models contribute to the certainty of whether domain components are mutually exclusive and collectively exhaustive. In relatively new domains, it might be difficult to use existing literature to establish a comprehensive set of domain components. In such cases, literature should only be used as a good theoretical starting point and other means to establish the domain components are needed. De Bruin *et al.* (2005) stated that literature, no matter how comprehensive, will probably not provide sufficient information to determine the level of detail needed for distinct capability areas. Therefore, it is recommended to use exploratory research methods to determine capability areas (Becker *et al.*, 2009; Lahrmann *et al.*, 2011). These methods include the Delphi technique, Nominal Group technique, case study interviews and focus groups.

4.2.3 Maturity levels

A sequence of maturity levels represents the typical evolutionary path of capability areas in a certain domain under consideration (Becker *et al.*, 2009). The bottom stage represents an initial state that is characterised by weak capabilities in the specific domain, where the highest stage represents total maturity with regards to the capability areas. Maturity levels are a number of cumulative stages where higher stages build on the requirements of lower stages (De Bruin *et al.*, 2005). This also correlates with Paulk *et al.* (1993) who stated that continuous process improvement is based on many small, evolutionary steps rather than on revolutionary innovations.

Different maturity models have a different number of maturity stages, but it is important that the different stages are distinct and well-defined (De Bruin *et al.*, 2005). There must also be a logical progression through the stages. Stages are named with short labels and stage definitions are developed. Stage names should give a clear indication of the intent of the stage. Stage definitions expand stage names and should give a summary of the significant requirements and characteristics of the stage.

Paulk *et al.* (1993) derived five maturity stages for continuous process improvement. The five levels are *Initial*, *Repeatable*, *Defined*, *Managed* and *Optimising*. According to De Bruin *et al.* (2005) these levels have wide practical use. Each preceding level is a necessary foundation for the next level and it would thus be counter-productive to skip levels (Paulk *et al.*, 1993).

Paulk *et al.* (1993) described the continuous process improvement of capability areas between the different stages. From the initial step capability areas progress to a disciplined process on the repeatable level, and then from repeatable to defined according to a standard, consistent process. From defined they progress to managed, by a predictable process. From managed they progress to optimising, which is a continuously improving process. This maturity progression is illustrated in Figure 4.1.

At the initial level, a stable environment for developing and maintaining processes is not provided. Planned procedures are often abandoned during crises. Objectives are often reached late and at high cost and success is dependent on individuals. Successful processes are not repeatable across the organisation (Paulk *et al.*, 1993).

A repeatable level is achieved when policies and procedures for managing processes are established. Basic process-management discipline enhances process capability and basic management controls are installed. Standards are defined and the processes of the organisation are disciplined, because they are stable and repeatable (Paulk *et al.*, 1993).

At the defined level, the processes across the organisation are documented. This includes software-engineering and management processes that are integrated into a coherent whole that is well-defined. A well-defined process includes readiness criteria, inputs, standards and procedures for performing the

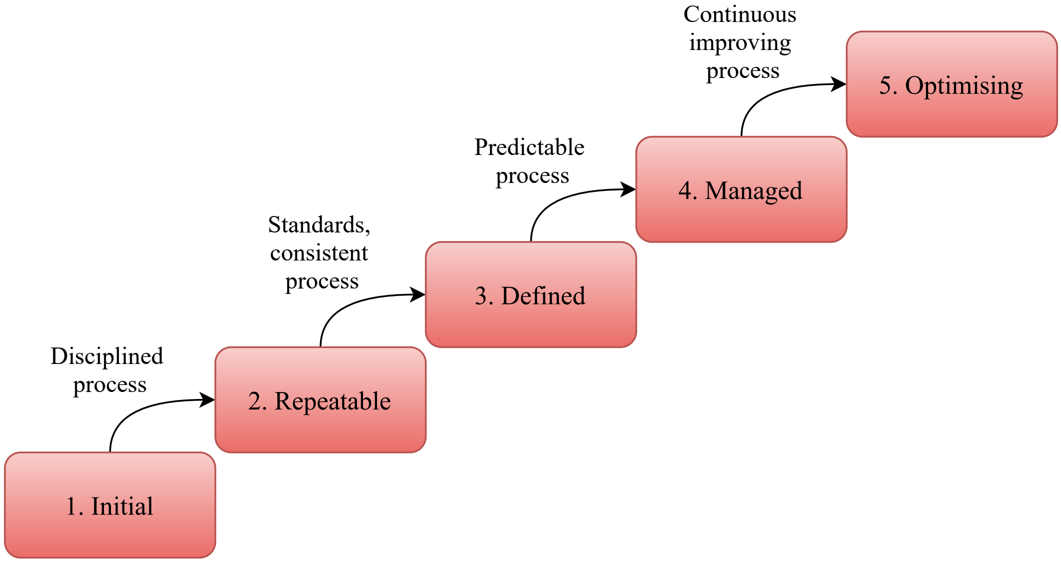


Figure 4.1: Capability maturity levels (Paulk *et al.*, 1993)

work, verification mechanisms, outputs and completion criteria. Level three processes are used to achieve effectiveness in the organisation. Processes are standardised and consistent. A common, organisation-wide understanding of the activities, roles and responsibilities exists (Paulk *et al.*, 1993).

At the managed level quantitative quality goals are set and measurements are well-defined and consistent, which ensures that processes operate within measurable limits. Predictable high quality is maintained. Available data is collected from defined processes and analysed organisation-wide. These measurements create the quantitative foundation for evaluating processes. Because processes are stable and measured, exceptional cases can be identified and addressed when they occur (Paulk *et al.*, 1993).

Continuous process improvement is the focus at the optimising level. Weaknesses in processes can be identified and the process can be improved by preventing defects. Changes to the process can be proposed through the use of data to perform cost-benefit analysis on new technologies. Innovations that exploit best process practices are identified and transferred successfully. Chronic waste from rework or random variation is reduced by changing the system to eliminate common causes of inefficiency. Technology and process improvements happen as normal business activities. Improvements are introduced both as incremental improvements on the existing process and as innovations in technology and methods (Paulk *et al.*, 1993).

4.3 Using a defined methodology to develop maturity models

It is very beneficial to use a defined methodology to develop maturity models to counter the countless criticism which maturity models have received in the past (Röglinger *et al.*, 2012). Some of these reasons include poor theoretical foundation and the impression of a falsified certainty to achieve success (Mettler, 2010*b*). Other criticism has been that maturity models oversimplify reality through step-by-step recipes that lack empirical foundation (De Bruin *et al.*, 2005). Maturity models also disregard that there might be multiple other maturation paths that have the same end result (Teo and King, 1997). Another criticism is that multiple similar maturity models exist (Röglinger *et al.*, 2012). Maturity models often have ambiguous results due to inadequate documentation on how to develop maturity models and the insufficient testing of the models (De Bruin *et al.*, 2005). Therefore, to mitigate criticism, a design process and a design product perspective are increasingly being used to develop maturity models (Röglinger *et al.*, 2012). Mettler (2010*b*) also synthesised very important design decisions into a generic maturity model development framework that consists of five design steps and eighteen decision parameters to help with the development of maturity models.

Although the thorough documentation of the development process of a maturity model is important to counter the criticism against maturity models, Mettler (2010*b*) did an extensive literature review that found that only a few developed maturity models disclose their design process and their underlying design decisions. Their literature search found only three well-defined maturity model design methodologies which include: (i) the design methodology according to De Bruin *et al.* (2005); (ii) the design methodology according to Becker *et al.* (2009); and (iii) the design methodology according to Mettler (2010*a*). These methodologies differ in the details of their model, but they agree substantially on common elements (Mettler, 2010*b*).

This section describes different maturity model design methodologies, design decisions and design principles that should be incorporated into the development of a maturity model. The incorporation of these components ensures that a sound methodology is followed that results in a usable maturity model. The methodologies included in this study are the ones according to De Bruin *et al.* (2005) (Section 4.3.1) and Becker *et al.* (2009) (Section 4.3.2). The two methodologies described in the studies of De Bruin *et al.* (2005) and Becker *et al.* (2009) were included in this study because they have great prominence in the field of maturity models, as both articles have more than 800 citations. Mettler (2010*b*) also stated that these studies are among the few studies that disclosed their design process, which is another reason why these design methodologies are described in the following sections. The design methodology according to Mettler (2010*a*) was excluded from this study as it has only 44 citations and does not share the prominence that De Bruin *et al.* (2005) and Becker *et al.* (2009) boast. Design decisions according to Mettler (2010*b*) are then described (Section 4.3.3). The design decisions of Mettler (2010*b*) were included in the study, as these design decisions were derived for five generic maturity model design process steps based on the work of De Bruin *et al.* (2005), Becker *et al.* (2009) and Mettler (2010*a*) and are therefore complementary to the studies of De Bruin *et al.* (2005) and Becker *et al.* (2009). Lastly, maturity model design principles are also discussed (Section 4.3.4).

4.3.1 Developing a maturity model according to De Bruin *et al.* (2005)

De Bruin *et al.* (2005) proposed a methodology that outlines the main phases of generic maturity model development. The phases are generic, but the order is important, because decisions of previous phases impact the phases that follow. Progression through some of the phases are iterative, for instance when it is necessary to revisit and adjust decisions made in earlier phases (De Bruin *et al.*, 2005). The main phases are (De Bruin *et al.*, 2005):

1. **Scope**

   When the scope of the model is determined, it sets the boundaries for the application and use of the model. This involves determining the focus of the model and identifying stakeholders to assist with model development. The focus of the model refers to the domain the maturity model will target. This determines the specificity and extensibility of the model and distinguishes it from other models (De Bruin *et al.*, 2005). Next, different stakeholders from academia, industry, non-profits and government are identified to assist in the development of the model. De Bruin *et al.* (2005) also suggested doing an extensive literature review of the specific domain, related domains and maturity models during the scoping phase. These give a good understanding of historical and present domain challenges.

2. **Design**

   During the design phase the design or architecture of the model is determined (De Bruin *et al.*, 2005). This forms the basis for further development and application. The criteria of the significant decisions of this phase are the audience, method of application, driver of application, respondents and application. Here, the needs of the intended audience are incorporated and how these needs will be met is determined. It is necessary to strike the right balance between model simplicity and complexity for the intended audience. During this phase the progression of maturity levels is determined, how many levels there should be and whether the levels are determined using a top-down or bottom-up approach. It should also be decided here whether to use a stage-gate approach that will determine the need for domain components and capability areas of the specified domain.

3. **Populate**

   The next phase is to populate the model architecture (De Bruin *et al.*, 2005). This is where the content of the model is determined. What needs to be measured and how it is measured are determined during this phase. It is crucial to identify domain components and capability areas for complex domains. For a mature domain, an extensive literature review can be used to determine domain components, but it is not very likely that literature will be sufficient to determine domain sub-components. Exploratory research methods are recommended to determine this layer of detail (De Bruin *et al.*, 2005). Further, it is also important to determine how maturity measurement can occur. This includes determining the instruments used to conduct the assessment and specifying the fitting questions and measures within the instrument. It is recommended to use a quantitative method like a survey (De Bruin *et al.*, 2005).

4. **Test**

   Phase four ensures the model is tested for relevance and rigour after it is populated (De Bruin *et al.*, 2005). The model construct and the model instruments need to be tested for validity, reliability and generalisability. Construct validity means face and content validity. Face validity scrutinises whether the population of the model was done by using tools like focus groups and interviews and whether the model can be considered as complete and accurate with regard to the identified scope of the model. Content validity assesses to what extent the domain is represented. A good measurement of content validity is the extent of the literature review and breadth of the domain covered. The assessment instruments should also be assessed for validity. This ensures they measure what it was intended to measure. Their reliability should also be assessed to ensure the results obtained are accurate and repeatable (De Bruin *et al.*, 2005).

5. **Deploy**

   The second last phase is the deployment phase where the model is made available for use and the generalisability of the model is verified.

6. **Maintain**

   Lastly, the relevance of the model will be ensured only by maintaining the model over time (De Bruin *et al.*, 2005). To establish the generalisability of the model successfully, it is essential to ensure that a high volume of the model can be applied. As the domain knowledge and model understanding expands, it will be necessary to update the model (De Bruin *et al.*, 2005).

4.3.2 Developing a maturity model according to Becker *et al.* (2009)

The procedure model of Becker *et al.* (2009) to develop maturity models used a scientific approach. The seven guidelines of Design Science (DS) defined by Hevner *et al.* (2004) were used as the basis of their argument. Venable and Baskerville (2012) defined Design Science Research (DSR) as:

> "Research that invents a new purposeful artefact to address a generalised type of problem and evaluates its utility for solving problems of that type."

Venable and Baskerville (2012) further explain four important concepts in their definition. Firstly, a purposeful artefact is any kind of artefact that achieves some human purpose, whether it is a product or a process and it can be a technology, a tool, a methodology, a technique, a procedure, a combination of these or any other means to achieve a purpose. Secondly, an invention means the creation, design, improvement or adaptation. Thirdly, addressing a generalised type of problem means that the artefact can be applied over and over again for various occurrences of that problem type. Fourth, evaluation proves that the developed artefact is useful to solve the problems it was created for. Through rigorous evaluation the utility and that the knowledge created is true and useful are proven. Venable and Baskerville (2012) claims that without evaluation, there is no science in DSR.

Becker *et al.* (2009) developed criteria for the development of maturity models from the DS guidelines defined by Hevner *et al.* (2004). Secondly, they used these criteria to serve as a basis to compare sparsely documented maturity models and thirdly, they generalised and consolidated the different models they studied to derive their generally applicable model. Their procedure model has eight phases for the development of maturity models. Based on the seven guidelines, the eight phases are:

1. **Problem definition**

   During the problem definition phase, both the targeted domain and the targeted group are important to be identified, as well as demonstrating the demand for the maturity model (Becker *et al.*, 2009).

2. **Comparison of existing maturity models**

   The next phase is the comparison of existing maturity models. The comprehensive comparison of existing maturity models is needed for a reasoned determination of the design strategy (Becker *et al.*, 2009).

3. **Determination of development strategy**

   The most important basic strategies stipulated are a completely new model design, the enhancement of an existing model, the combination of several models into a new one, and the transfer of structures and contents from existing models to new applications (Becker *et al.*, 2009).

4. **Iterative maturity model development**

   The iterative maturity model development phase is the most important phase of the procedure model of Becker *et al.* (2009). The sub-steps that are iterated are selecting the design level, selecting the approach, designing the model section and testing the results. The fundamental structure of the maturity model is defined by the highest degree of abstraction. This provides the architecture of the maturity model. From there, the model is further developed by devising individual domain components and their capability areas that fit into the model architecture. Appropriate methods must be chosen to develop each abstraction level of the model. This includes, but is not limited to, the use of literature analysis and exploratory research methods like the Delphi method and creative techniques. The results of the developed part must be tested for comprehensiveness, consistency and problem adequacy.

5. **Conception of transfer and evaluation**

   This phase determines how the results will be transferred to academic and user communities (Becker *et al.*, 2009). The different possibilities for the evaluation of the problem solution that the maturity model proposes should be included in the transfer design.

6. **Implementation of transfer media**

   Next, the implementation of the transfer media phase ensures that the maturity model is accessible to all defined user groups in the way it was planned (Becker *et al.*, 2009).

7. **Evaluation**

   The evaluation phase is important for comparing the defined goals of the model with the real-life observations. It establishes whether the projected benefits and an improved solution for the defined problem is reached. Case studies have commonly been used. Based on the outcome of the evaluation process, a reiteration of the design process might be needed (Becker *et al.*, 2009).

8. **Rejection of maturity model**

   If the evaluation reveals negative results, it may lead to a rejection of the model (Becker *et al.*, 2009).

4.3.3 Using design decisions for developing maturity models

Mettler (2010*b*) derived five generic maturity model design process steps with a number of design decisions specified for every step, based on the comparison of the maturity model development processes specified by De Bruin *et al.* (2005), Becker *et al.* (2009) and Mettler (2010*a*). As the previous design methodologies are largely generic, they do not aid developers with essential design decisions (Mettler, 2010*b*). When crucial design decisions are neglected, it can have a critical impact on the research product and its adoption. Therefore, Mettler (2010*b*) extended the five generic maturity model design steps with 18 crucial decision parameters that are illustrated in Table 4.1. The different steps are (Mettler, 2010*b*):

1. **Identify need or new opportunity**

   The two decision parameters in this step are deciding on the novelty of the topic and the innovation of the maturity model (Mettler, 2010*b*). Novelty is important to determine whether there is a need for an explanation in practice and whether there are some use cases to underpin theoretical assumptions on the maturity model. The second decision parameter is to decide whether the model is completely new, a variant or a version of an existing model. This has a significant effect on subsequent design decisions.

2. Define scope

   According to Mettler (2010*b*), during this step the most important design decisions are made. The first decision is to determine the breadth of the study, and whether a general issue or a subject-matter issue is being addressed. Secondly, the depth parameter determines the altitude that the maturity model operates on (Mettler, 2010*b*). This parameter is characterised by either the group level, organisation level, inter-organisational level or on a global or societal level. Lastly, the potential audience of the proposed model must be taken into consideration for development. It should be considered whether the model is developed for the need of management-oriented or technology-oriented audiences, or both.

3. **Design model**

   During this step the actual maturity model is built (Mettler, 2010*b*). Therefore, it is essential to determine what is understood by the term 'maturity'. The three prevalent different concepts of maturity are process-focused, object-focused and people-focused. A process-focused understanding implies centring on activities and work practices that define more effective procedures. Object-focused understanding focuses on the features of work products in order to improve their mode of operation. Lastly, people-focused means the emphasis on soft capabilities. The second decision to make is whether the goal function of maturity focuses on one dimension or on multiple dimensions. The goal function is the way in which maturity is progressed. The concept of maturity tentatively determines the goal function. For instance, process-focused almost always implies efficiency. But a maturity model can target multiple goal functions. Next, the nature of the design process is determined. This identifies the knowledge base for deriving the maturity levels, the metrics and the corresponding improvement recommendations. This decision affects the research methods that are used (such as literature reviews versus focus group discussions). The next three decision parameters are the design product, application method and respondents. The design product is important for the determination of the development team. This has an effect on the choice of application method (self or third-party assessment or outsourced) and the respondents for data collection.

4. **Evaluate design**

   During this step it is important to decide how the model is verified and validated (Mettler, 2010*b*). It is possible to evaluate the way the model was constructed or to evaluate the model itself. It is recommended that both should be evaluated (Mettler, 2010*b*). After this, the point in time of the evaluation should be decided on. Lastly, the evaluation method parameter is comprised of deciding between using artificial or naturalistic methods (Mettler, 2010*b*).

5. **Reflect evolution**

   During this step the design mutability is decided (Mettler, 2010*b*). Alterations are needed to be made, because the subject under study is growing. Becker *et al.* (2009) stated that maturity models inherently become obsolete. This is due to new scientific insights, technological progress or changing conditions. Changes in form and functioning ensure the standardisation and global acceptance of the model. Also, the frequency (non-recurring versus continuous) and the structure of change (modifications made exclusively by the developer or openly by other users) needs to be determined (Mettler, 2010*b*). Lastly, it should be decided whether the model itself is disseminated to an exclusive group of people and organisations or whether it is freely available (Mettler, 2010*b*).

Table 4.1: Design decisions (Mettler, 2010*b*)

Design activity	**Design parameter**	**Characteristics**			
Identify need	Novelty	Emerging	Pacing	Disruptive	Mature
	Innovation	New	Variant	Version	
Define scope	Breadth	General issue		Specific issue	
	Depth	Individual/ Group	Organisation	Inter-organisational	Global/ Society
	Audience	Management oriented	Technology oriented	Both	
Design model	Maturity concept	Process-focused	Object-focused	People focused	Combination
	Goal function	One-dimensional		Multi-dimensional	
	Design process	Theory-driven	Practitioner based	Combination	
	Design product	Textual description of form	Textual description of form and functioning	Instantiation (software)	Combination
	Application method	Self-assessment	Third-party assisted	Certified professionals	
	Respondents	Management	Staff	Business partners	Combination
Evaluate design	Subject of evaluation	Design process	Design product	Both	
	Point of time	Ex-ante	Ex-post	Both	
	Evaluation method	Naturalistic	Artificial	Combination	
Reflect evolution	Subject of change	None	Form	Functioning	Form and functioning
	Frequency	Non-recurring		Continuous	
	Structure of change	External/ open		Internal/ exclusive	
	Dissemination	Open		Exclusive	

4.3.4 Design principles for maturity models

According to Pöppelbuß and Röglinger (2011), extensive research has been done on the design process and structural components of maturity models, but a holistic understanding of the design principles that a maturity model should meet was lacking. Therefore, Pöppelbuß and Röglinger (2011) developed a framework of general design principles of form and function that maturity models should meet.

Pöppelbuß and Röglinger (2011) divided their design principles into three categories. These categories are: (i) basic principles; (ii) principles for a descriptive purpose of use; and (iii) principles for a prescriptive purpose of use. The application-specific purposes of use of maturity models are distinguished between descriptive, prescriptive and comparative. Pöppelbuß and Röglinger (2011) deliberately omitted the comparative purpose of use in their design principles framework.

A descriptive purpose of use maturity model is applied to assess the as-is state of the current capabilities with regard to given criteria (De Bruin *et al.*, 2005). A descriptive purpose of use maturity model is used as a diagnostic tool.

Prescriptive purpose of use maturity models identify desirable maturity levels and provide guidelines on improvement measures (De Bruin *et al.*, 2005). This means that specific and detailed courses of action are suggested.

A maturity model serves a comparative purpose of use when it permits internal or external benchmarking (Pöppelbuß and Röglinger, 2011). Maturity levels of similar business units and organisations are compared when sufficient historical data is available for assessment (De Bruin *et al.*, 2005).

The groups of design principles that Pöppelbuß and Röglinger (2011) developed are nested (Röglinger *et al.*, 2012). This means that descriptive models should adhere to the basic principles as well, and not only the design principles for descriptive purpose of use (Pöppelbuß and Röglinger, 2011). Prescriptive models should adhere to basic, descriptive and prescriptive principles. The basic design principles should be addressed independently of a specific purpose of use. The different principles are discussed further. Basic design principles are discussed in Section 4.3.4.1, descriptive purpose of use principles are discussed in Section 4.3.4.2 and prescriptive purpose of use principles are discussed in Section 4.3.4.3.

4.3.4.1 Basic design principles

There are four basic design principles (Pöppelbuß and Röglinger, 2011). These four design principles are: (i) providing basic information; (ii) defining central constructs related to maturity and maturation; (iii) defining central constructs related to the application domain; and (iv) documenting the basic information, central constructs and their interrelations in a target-group oriented manner.

Basic information that needs to be included are the application domain and prerequisites for applicability, the purpose of use, the target group, the class of entities under investigation, how the maturity model is different from other related maturity models. The design process and how the model was validated should also be documented and communicated in a understandable way for the target group (Pöppelbuß and Röglinger, 2011).

The different central constructs related to maturity and maturation that should be defined are maturity and dimensions of maturity, maturity levels and maturation paths, the available levels of granularity of maturation and the theoretical foundations of evolution and change (Pöppelbuß and Röglinger, 2011). Defining maturity means that the concept of maturity must be defined in relation to the class of entities and application domain under investigation. Maturity can be one-dimensional where it focuses on one component like object or process maturity. It can also be multi-dimensional. A multi-dimensional approach facilitates the definition of assessment criteria for a descriptive purpose of use, as well as the classification of improvement measures for a prescriptive purpose of use. Maturation paths consist of maturity levels that are identified by a concise descriptor. The maturation rationale with regard to the logical relationship between successive levels needs to be communicated. The levels of granularity of maturation can either be a high level of abstraction or a low level of abstraction. These different levels of granularity can be established by structuring maturity models hierarchically into multiple layers. Lastly, the underpinning theoretical foundations should be explicated in terms of the way change usually happens in the specific domain (Pöppelbuß and Röglinger, 2011).

4.3.4.2 A descriptive purpose of use

The two design principles for a descriptive purpose of use are: (i) proposing assessment criteria for each maturity level and available level of granularity that exhibits a high level of intersubjective verifiability; and (ii) including a target group-oriented assessment methodology that is also intersubjectively verifiable (Pöppelbuß and Röglinger, 2011). Intersubjective verifiability of criteria ensures the comparability of the maturity assessments. Intersubjective verifiability of criteria is reached by having precise and concise descriptions of corresponding levels and it should be clear to discriminate between levels (Pöppelbuß and Röglinger, 2011).

The target group-oriented assessment methodology needs to include a procedure model that guides model users through maturity assessments, advice on the assessment of criteria, advice on the adaptation and configuration of criteria according to different situational characteristics and available knowledge from previous applications (Pöppelbuß and Röglinger, 2011). The procedure model should elaborate on the assessment steps, their interplay and specifically how to obtain the values of the criteria. The results from an assessment also need to be correct, accurate and repeatable (Pöppelbuß and Röglinger, 2011).

4.3.4.3 A prescriptive purpose of use

A prescriptive purpose of use maturity model has three additional design principles (Pöppelbuß and Röglinger, 2011). The three design principles are: (i) including improvement measures for each maturity level and level of granularity; (ii) including a decision calculus for selecting improvement measures; and (iii) including a target group-oriented decision methodology (Pöppelbuß and Röglinger, 2011).

A decision calculus enables model users to choose improvement measures (Pöppelbuß and Röglinger, 2011). The decision calculus helps evaluate different alternatives with respect to objectives. This helps to identify which alternative reaches the objective best. The relevant objectives and the relevant factors that influence performance should be explicated. The decision calculus should also discriminate between external reporting and an internal improvement perspective (Pöppelbuß and Röglinger, 2011).

Maturity models for a prescriptive purpose of use should also define a target group-oriented decision methodology (Pöppelbuß and Röglinger, 2011). The procedure model that guides model users through the steps of improvement measure selection is the most essential component. Advice on the assessment of variables should be provided. Advice on how to concretise and adapt improvement measures should be given by the decision methodology. It should also give advice on how to adapt and configure the decision calculus. Lastly, available knowledge from previous applications should be reported (Pöppelbuß and Röglinger, 2011).

4.4 Maturity models in the healthcare data management domain

In this section maturity models are reviewed in the data management domain in the healthcare sector. This was done to gain a better understanding of the different challenges that are persistent in this domain and how maturity models have strived to find solutions for these challenges. This section also reviews what can be learnt from previous models by determining how these models were developed, what components they addressed, and identifying the maturity concepts and goal functions they specified. The scoping of the relevant maturity models also enables the identification of what is currently still missing in the design of current models and how they fail to fully address the challenges they set out to do. This enables the incorporation of components into a new model to ascertain that a new model achieves its goal to assist in addressing the healthcare data management challenges. These insights were considered when the new maturity model was developed as described in Chapter 5. The sections to follow consist of an explanation of the healthcare data management literature review process (Section 4.4.1) and an exposition of the contributions from existing healthcare data management maturity models (Section 4.4.2).

4.4.1 Literature review process of healthcare data management maturity models

As part of the scoping phase of the development of a maturity model, De Bruin *et al.* (2005) suggested doing an extensive literature review on the domain of interest and other relevant domains to gain understanding of the historic and persistent present challenges in the domain. Tarhan *et al.* (2020) conducted a multi-vocal literature review on maturity assessment and maturity models in healthcare. The multi-vocal review included grey literature and yielded 101 sources.

Tarhan *et al.* (2020) concluded that researchers should check the list of maturity models and the subject focuses that they reported before proposing a new model. This eliminates the possibility of creating identical studies and enables the chance to adapt or use existing maturity models. Not all the sources that Tarhan *et al.* (2020) found in their study were reviewed for this study. Only the literature relevant to this study was reviewed. This is literature that focused specifically on data management in healthcare and not merely healthcare in general. Tarhan *et al.* (2020) categorised the different sources into different categories. The different categories were used to determine which literature was relevant.

A set of exclusion criteria was used to determine which studies were relevant to this study. Firstly, studies were excluded according to their application scope. Application scope categories that were excluded were "single department", "multiple departments" and "single hospitals" as this study focuses on healthcare data management on a broader scale than single hospitals. Not all sources were categorised into any of the application scope categories and those that were not categorised were excluded. This yielded 85 remaining studies.

The next exclusion criterion was according to the type of paper (contribution facet). Here, only studies that focused on the presentation of new maturity models were not excluded as this study strives to propose a new model too. This yielded 54 studies.

Technological aspects covered was the next exclusion criterion. The studies that focused on information systems (IS) and information technology (IT) were not excluded from the study. This study focuses on the data management of healthcare and therefore these components were important to be reviewed. This yielded 44 articles.

Of the remaining articles, the titles and the focus area of the maturity model were inspected. Literature that focused on the management of data in healthcare was not excluded from the study. These included papers on data management security and privacy, interoperability, storage, analysis, etc. Anything that had to do with the data management of healthcare was not excluded. This yielded 30 articles.

Two more articles were included from the references of the articles as they gave valuable insights to healthcare data management. Therefore, the eventual number of articles relevant to this study was 32 articles.

These were the papers included for the study as they were relevant to the topic of healthcare data management. These were read through as De Bruin *et al.* (2005) suggested to gain valuable information on the challenges that past maturity models tried to address, which aspects of data management the other models covered, what can be learnt from the development process of previous healthcare maturity models and how this model can be distinguished from them. The studies focused on healthcare data management indicate the usefulness of maturity models in the domain of healthcare data management, as my studies developed maturity models to address healthcare data management challenges.

The extensive review of studies done in the domain of healthcare data management depicts there is really a need for a new maturity model in this domain. There is a need for a new maturity model, because no existing maturity model was identified that covers the whole scope of healthcare data management challenges across the healthcare data management value chain. Therefore, it is necessary to develop a maturity model that assists in addressing the healthcare data management scope of challenges.

4.4.2 Contributions from existing healthcare data management maturity models

It was found that there are two main approaches when developing healthcare data management maturity models (Carvalho *et al.*, 2016). On the one hand, highly specialised models that resulted in a health sub-system were developed and on the other, comprehensive models that represent healthcare data management as a whole, were developed. These specialised and comprehensive maturity models that focus on data management in healthcare are discussed in this section. The studies are categorised in the following sections according to their focus area. Firstly, existing maturity models that had a general focus are described (Section 4.4.2.1), followed by a description of digital healthcare data maturity models (Section 4.4.2.2), a description of healthcare interoperability, networkability and cooperation maturity models (Section 4.4.2.3), a description of healthcare data security maturity models (Section 4.4.2.4), a description of healthcare data analysis maturity models (Section 4.4.2.5), and a description of the assessment of information technologies in hospitals (Section 4.4.2.6). Lastly, the knowledge transferred from these existing maturity models are summarised in tabular form in Section 4.4.2.7.

4.4.2.1 General healthcare maturity models

Flott *et al.* (2016) proposed a patient-centred framework for the evaluation of the digital maturity of health services. It measures digital maturity in a multi-dimensional way and the five themes they focused on are: (i) general evaluation methodology; (ii) resources and ability; (iii) usage; (iv) interoperability; and (v) impact. The maturity levels progress from paper-based, to basic electronic record-keeping, clinical processes, advanced disease management, integrated care and population-impact. The first four maturity levels progress serially and the last two are iterative (Flott *et al.*, 2016).

Liu *et al.* (2011) did a nation-wide investigation on the e-health maturity in Taiwan. They proposed a multi-dimensional maturity assessment model. Technology usage is not sufficient to evaluate the IT level of an organisation and therefore, Liu *et al.* (2011) proposed a six-dimensional model to assess e-healthcare maturity. The six dimensions included IT infrastructure, IS staff, user awareness, IS planning and control, integration and application portfolio (Liu *et al.*, 2011).

Sharma (2008) proposed a 7-level maturity model that exhibits how healthcare processes can reach maturity up to a national level. Sharma (2008) included all associated service providers in the healthcare process that are adaptable to any provider at any level of maturity. The seven levels are the hospital administration level, the hospital enterprise level, the Electronic Medical Record (EMR) basic level, the clinical decision support level, the clinical research level, the regional level and the national level. At each level the entities, departments and infrastructure involved are included.

Vidal Carvalho *et al.* (2019), Carvalho *et al.* (2019*a*) and Carvalho *et al.* (2019*b*) conducted three studies to develop a maturity model of hospital information systems. For the first study a laborious method was used to identify the most important influencing factors of hospital information systems (Vidal Carvalho *et al.*, 2019). The six most important influencing factors were: (i) data analysis; (ii) strategy; (iii) people; (iv) EMRs; (v) information security; and (vi) systems and IT infrastructure. In the second study, the Hospital Information System Maturity Model (HISMM) was developed with these six most important influencing factors (Carvalho *et al.*, 2019*a*). Lastly, HSIMM was extended with another component that focuses on the maturity of healthcare data analytics (Carvalho *et al.*, 2019*b*).

4.4.2.2 Digital healthcare data maturity models

Johnston (2017) did a digital maturity assessment that assesses the state of readiness of the organisation to integrate digital technologies. To do this, the capability and compatibility of the information system of the organisation to communicate and interface within the organisation and across organisations are assessed (Johnston, 2017). They believe that electronic medical records are the cornerstone to the advancement of digital maturity in healthcare. This enables the management and transfer of digital patient data across clinical and geographical boundaries that spans across the entire network of care providers. The system-wide approach to integrate new technologies to establish new communication lines is necessary to enable this (Johnston, 2017). Johnston (2017) focused on providing patient-centred services that meet local needs within a national framework.

The Informatics Capability Maturity Model (ICMM) that assesses healthcare maturity out of an informatics perspective was proposed by Directorate Informatics . Health Informatics is the required knowledge, skills and tools needed to get the right information to the right person at the right time to promote healthcare and promote health (Directorate Informatics). The five capability dimensions they looked at are managing information, using business intelligence, using information technology, aligning business and information and managing change. These five dimensions were measured against the maturity scale of five levels that are (Directorate Informatics): basic, controlled, standardised, optimised and innovative. They also listed an extensive list of benefits across the different capability dimensions.

Rimmer *et al.* (2014) developed a maturity model that measures the maturity of use of electronic medical records. They looked at maturity as the escalation of effectiveness and efficiency. The five maturity levels specified for this model are front office administration, EMR basics, full EMR, proactive care / data-driven practice and community-shared care.

Grindle *et al.* (2013) stated the changes that the combination of cloud computing, mobility and data analytics could bring to healthcare. They stated five benefits that can be reaped through this combination. Firstly, it can create the agility for health providers and insurers to flex business models, collaborate at low cost and high speed, meet changing regulations and deliver better patient care. Secondly, it can help ensure seamless, personalised healthcare through the sharing of secure and ubiquitous data. It can also shift the location of chronic healthcare from the hospital or clinic to the home that enables better patient care and comfort at lower cost. Health and non-health big data can be harnessed and analysed in the cloud that can improve public health through preventative well-being monitoring. The last benefit they mention is that access to health services can be transformed without heavy investments in physical hospitals and clinical centres. They proposed a very basic four-level cloud maturity model for the healthcare industry (Grindle *et al.*, 2013). The four main phases are experimentation, foundation building, optimisation and innovation.

4.4.2.3 Healthcare interoperability, networkability and co-operation maturity models

Naudet and Chen (2012) conducted a study to evaluate eHealth interoperability using the Maturity Model for Enterprise Interoperability (MMEI). Interoperability is the ability of two or more systems or components to exchange information and to use the information that was exchanged (Naudet and Chen, 2012). The MMEI defined the three aspects of interoperability as conceptual, technical and organisational interoperability. Conceptual interoperability ensures that the exact meaning of exchanged information is understandable by all other systems involved. Technical interoperability is concerned with linking systems and services and organisational interoperability defines the responsibilities and authority so that interoperability takes place under a good environment (Naudet and Chen, 2012). The five maturity levels are unprepared, defined, aligned, organised and adaptive.

Venesco (2015) also conducted a study to improve interoperability. To ensure the widespread use of electronic medical records and electronic exchange of information amongst various data users, the accuracy of each person's identity is important and the frequency of errors in matching patient records needs to be reduced (Venesco, 2015). They proposed a data management maturity model for individual data matching that consists of three aspects, namely: (i) data quality; (ii) processes; and (iii) relevant regulations. These aspects are measure against a five-level maturity scale.

There are a number of studies that focus on maturity models of the cooperability and networkability within and between hospitals (Fitterer and Rohner, 2010; Blondiau *et al.*, 2013; Mettler and Blondiau, 2012). Networkability is the organisation's capacity to efficiently and rapidly engage relationships with business partners (Blondiau *et al.*, 2013) or in other words, networkability is both the internal and external capability of organisations to cooperate at both the business processes and underlying Information and Communications Technology (ICT) infrastructure levels (Fitterer and Rohner, 2010). It is not limited to system-to-system interoperability on a software application level (Fitterer and Rohner, 2010). Instead, the three layers of corporate interoperability defined by Winter (2003) are incorporated to measure the cooperation within and between hospitals comprehensively (Fitterer and Rohner, 2010; Blondiau *et al.*, 2013; Mettler and Blondiau, 2012). The strategic layer measures the ability of the hospital to cooperate with external partners, the organisational layer measures the ability to cooperate within the hospital and the information layer measures the technical capabilities of a hospital through the required infrastructure for the effective and efficient internal and external cooperation (Mettler and Blondiau, 2012; Blondiau *et al.*, 2013; Fitterer and Rohner, 2010). The four scales used by Mettler and Blondiau (2012) and Blondiau *et al.* (2013) were initial / ad hoc, committed, established / focused and optimised. Fitterer and Rohner (2010) used the five-point Capability Maturity

Model Integration (CMMI) set of maturity levels which are *initial*, *managed*, *defined*, *quantitatively managed* and *optimising*.

4.4.2.4 Healthcare data security maturity models

The Healthcare Information Security Adoption Model (HISAM) was developed by Impact Advisors (2015) that helps healthcare organisations measure their level of security. It describes seven levels of security preparedness across three serial categories which are awareness, technical and behavioural. The first two levels, that fall under the awareness category, are first steps and getting organised. The three technical levels are basic technology, standards and procedures, and advanced technology. The last two levels comprise the behavioural level and consist of educated enterprise and integrated security (Impact Advisors, 2015).

Frost & Sullivan (2015) also developed a maturity assessment model that assesses the maturity of the components needed against security breaches. Healthcare data is subject to identity theft and therefore, its security is very important. Frost & Sullivan (2015) stated many different types of data breaches and to address the different types, they split the different security components into three groups which are client, network and service groups. The Healthcare Security Maturity Model has three security levels which are baseline, enhanced and advanced (Frost & Sullivan, 2015).

4.4.2.5 Healthcare data analysis maturity models

Carvalho *et al.* (2019*a*) expanded the HISMM to incorporate Data Analytics (DA), resulting in the HISMM-DA. DA helps organisations be more productive and profitable when managing decision-making processes (Carvalho *et al.*, 2019*b*). The HISMM-DA adopted the six maturity stages of the HISMM. These maturity stages are adhocracy, starting the foundations, centralised dictatorship, democratic cooperation, entrepreneurial opportunity and integrated relationships.

Sanders *et al.* (2013) developed the Healthcare Analytics Adoption Model (HAAM) that focuses on enabling effective data analysis. Sanders *et al.* (2013) identified the trend of healthcare computerisation in three phases. These phases are data collection, data sharing and data analysis. Healthcare data management is entering phase three, but in order to enable effective data analysis, the first two phases are essential. In phase one computerisation specifically supports transaction-based workflow and data collection. This is characterised by widespread EMR adoption. In phase two, data is shared among members of the workflow team and is characterised by health information exchange (HIE). Phase three of computerisation involves the analysis of the collected and shared data which HAAM focuses on (Sanders *et al.*, 2013). They proposed nine levels that illustrate the progressive capability improvement of healthcare data analysis (Sanders *et al.*, 2013). According to Sanders *et al.* (2013) HIE has an overwhelming failure rate and they also believe that the deployment of EMR does not significantly improve healthcare quality or cost. Despite this, they believe that the deployment of EMR is fundamentally required to achieve the value that is accessible in healthcare through analytics (Sanders *et al.*, 2013).

4.4.2.6 The assessment of information technology capacities in hospitals

Jaana *et al.* (2009) conducted a study to present an instrument to assess IT capacities in hospitals. IT capacities in hospitals were categorised in three dimensions according to the model that Paré and Sicotte (2001) developed, namely functional IT capacities, technological IT capacities, and internal and external integration IT capacities. This means the instrument measures the extent of implementation of computerised processes, the extent of implementation of technological devices and thirdly, it measures the integration of internal administrative and clinical information and the extent of sharing of information with other external entities (Jaana *et al.*, 2009). The different computerised processes are further categorised under clinical applications, ancillary applications, patient management applications and administrative applications. The technological devices are further categorised under clinical technologies and administrative technologies (Jaana *et al.*, 2009).

4.4.2.7 Existing healthcare data management maturity model knowledge transfer

The insights gained from the literature study on healthcare data management maturity models were categorised according to the components of the derived value chain model from Section 3.1.1. The different healthcare data management maturity models focused on different components of data management. Some addressed various components and others focused on one specific component. What was learnt from the different studies under each value chain component is presented in tabular form in Appendix C. Knowledge from the existing maturity models was transferred with regard to: (i) the development of maturity models; (ii) healthcare data collection; (iii); data storage; (iv) data sharing; (v) data analysis; (vi) data usage; (vii) data privacy and security; (viii) data governance; (ix) data technology and infrastructure; (x) human contributions; and (xi) technological investments.

4.5 Conclusion on maturity models

During this chapter maturity models were introduced. The origin, purpose and value of maturity models were described. The basic structure of a working maturity model was identified. The necessary components included a specified domain, domain components, capability areas, maturity levels, and maturity descriptions. An argument to state the importance of using a defined methodology with design decisions and principles to develop maturity models was presented. Some examples of these defined methodologies were presented and the design decisions and principle were listed. Maturity models in the healthcare data management domain were also presented to indicate what has already been done in this domain, what challenges they strived to address and what could be learnt from these models.

This chapter identified maturity models as a suitable research product to be used to assist in addressing the healthcare data management challenges. The usefulness of maturity models in this domain is evident and maturity models have a long history of improving systems progressively. The healthcare data management challenges are not a problem that can be solved instantaneously and they should be addressed continuously and progressively. Maturity models can help improve healthcare data management; however, to ensure their usefulness, it is important to use a well-defined design methodology, with design decisions and design principles when developing maturity models. Therefore, to develop a maturity model in the context of the stated problem, a thorough design methodology should be followed. Also, the insights gleaned from existing maturity models in the context of healthcare data management should be used to develop an appropriate model. The proposed model should also be distinguishable from existing models to ensure that it contributes to the existing body of knowledge.

Chapter 5

Development of the healthcare data management maturity model

In this chapter, the development process of the Healthcare Data Management Maturity Model (HCDMMM) that aims to facilitate the process of assessing healthcare data management maturity is presented. The specific domain that the HCDMMM is developed for is the health care data management domain of developing countries. The literature from the previous chapters is used to support the development process of the HCDMMM and was used to assist in the development of the different components of the HCDMMM. Furthermore, literature from Chapter 4 is used to determine a defined development process with design decisions and design principles. Chapter 4 literature also consists of previous studies on maturity models on data management in the health sector. Knowledge from these studies was used to identify the requirements for a maturity model in the domain of healthcare data management and also to discover what insights could be gained from these previous studies. Input from Chapter 3 included the requirements that were considered during the development process. The data value chain with its different components was also incorporated in the development process to further define the different components of the HCDMMM. The first section of this chapter (Section 5.1), consists of the development methodology that was derived using the inputs from previous studies on maturity models. Subsequently, the identification of the need for another maturity model is described (Section 5.2). The scope that determines the specificity and extensibility of the developed maturity model is defined next in Section 5.3, and the design and population of the HCDMMM is described in Section 5.4. This is an iterative development process and the development of the HCDMMM is described per each individual component. The HCDMMM was evaluated through verification and validation as described in Chapter 6. In Section 5.5 the evaluated HCDMMM is presented: first, conceptually and then from a practical perspective as it will be transferred to end users. The reflection on future upgrades is discussed to ensure that HCDMMM does not become obsolete over time, but remains relevant (Section 5.6). Lastly, this chapter on the development and presentation of the HCDMMM is concluded (Section 5.7).

5.1 Healthcare data management development methodology

To develop the HCDMMM that will contribute to solving the healthcare data management challenges, a defined development methodology was used. The defined development methodology was derived from inputs from existing methodologies such as that of De Bruin *et al.* (2005) and Becker *et al.* (2009), the design decisions of Mettler (2010*b*) and the design principles of Pöppelbuß and Röglinger (2011). These defined methodologies, design decisions and principles are discussed in Section 4.3. The first step of the methodology was to identify the need for a new maturity model. This is to prevent the development of a maturity model in an already saturated domain and also to avoid the HCDMMM being similar to already developed models. Secondly, the scope of the study should be determined in order to establish clear boundaries of what the HCDMMM addresses. Different scoping decisions are made to help determine the scope. After that, the HCDMMM was designed and developed. The first step of the HCDMMM design was to make the necessary design and development decisions, such as deciding on the maturity concept, the goal function, design process, design product, the application method and the respondents. Subsequently, the design and development of the HCDMMM commenced. The design and development process comprises of development and evaluation steps that were iterated. The development steps consist of constructing the architecture of the HCDMMM and populating it. Following this, the developed model structure and content were verified. SMEs were consulted during this verification step and semi-structured interviews were held with them to assess the HCDMMM. If the verification step showed the need for refinements, the HCDMMM underwent another development iteration. After several iterations, when the HCDMMM structure and content was completed, it was subjected to a final verification step where healthcare data management SMEs, who were able to assess the HCDMMM comprehensively, responded by means of a questionnaire, as to whether the HCDMMM satisfactorily met the specified requirements. The verification step after the iterative model development process yielded some final refinements to the HCDMMM content and resulted in the development of the transfer media that ensured all the specified requirements were met satisfactorily. Figure 5.1 illustrates the defined design methodology that was used to develop the healthcare data management maturity model.

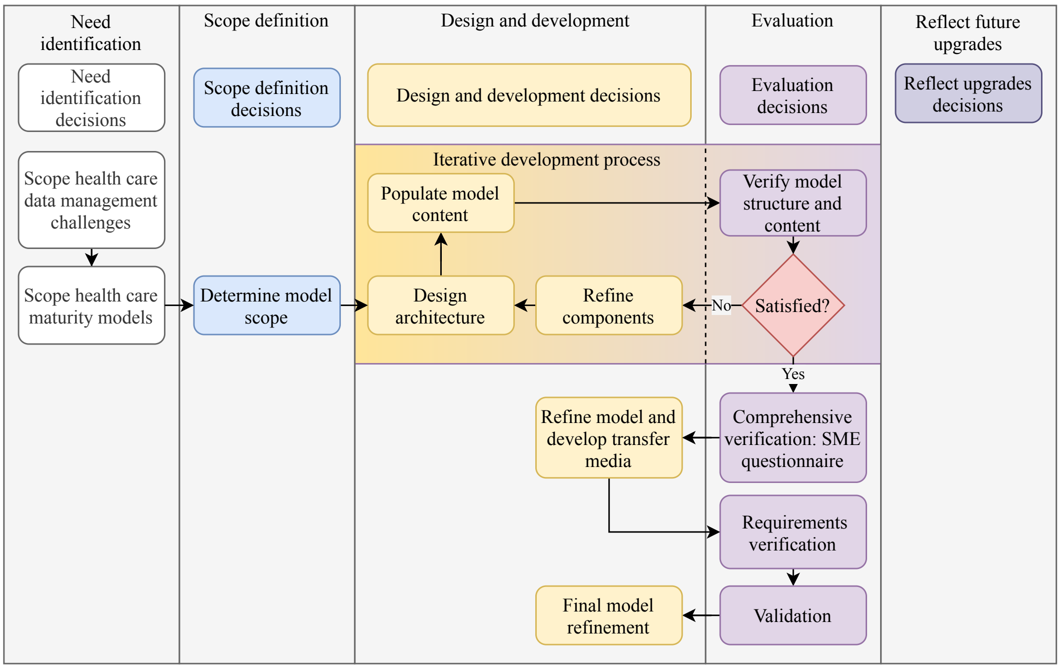


Figure 5.1: Model design methodology

5.2 Need identification

Becker *et al.* (2009) stated that before a maturity model is developed, the requirement for such a maturity model should be demonstrated. The scope of healthcare data management challenges presented in Section 2.3 indicate the need of a research product that will contribute to addressing the challenges in healthcare data management. During the multi-vocal literature study that Tarhan *et al.* (2020) conducted on maturity models in healthcare as discussed in Section 4.4.1, it was found that no existing maturity model adequately addressed the scope of healthcare data management challenges in developing countries and therefore, a need for a maturity model that will contribute to addressing the scope of healthcare data management challenges in developing countries arose. Developing a maturity model that contributes to addressing the challenges found in Section 2.3 ensures that the need for a maturity model that focuses on healthcare data management in developing countries, is justified. Tarhan *et al.* (2020) also stated that due to the increasing level of digitalisation of healthcare services, well-defined guidelines and success reports to rationalise and effectively manage the transition process are needed. It is thus argued that maturity models can contribute to addressing this need in the healthcare data management domain.

Reviewing literature extensively provides a sound understanding of historical and contemporary challenges in the domain (De Bruin *et al.*, 2005; Becker *et al.*, 2009). Existing literature that was reviewed in the domain of healthcare data management can be viewed in Section 4.4. The review of literature ensures that the proposed model does not produce an identical study to previous studies and that it contributes to the existing body of knowledge by addressing the challenges that other models have not yet addressed, or have not addressed adequately. Reviewing previous studies also contributes to assisting the development of the proposed model by adapting or using some of its components.

The design decisions of Mettler (2010*b*) that correlate with this section are novelty and innovation. With regard to novelty, data management is a 'pacing' field of study (as opposed to emerging, disruptive or mature). This means that data management practices are well established across many domains, but improvements to data management practices are still being gradually introduced. There are also emerging or disruptive components to data management, such as big data, which makes it a dynamic field of study. As healthcare data management is a 'pacing' field of study with well-established practices, it requires less explanation than an emerging or disruptive field, although emerging concepts in data management need to be explained in more detail. The second decision that was made was that a new model was going to be developed (as opposed to a variant or version of existing models). Although the proposed HCDMMM drew valuable insights from existing knowledge, the HCDMMM focuses specifically on addressing healthcare data management challenges in developing countries, as previous studies have not sufficiently addressed the scope of the problem statement of this study. The existing studies only investigated certain components of data management and do not address it holistically or on multiple system levels. The decisions and decision descriptions can be seen in Table 5.1.

Table 5.1: Need identification decisions (Mettler, 2010*b*)

Design parameter	**Chosen characteristic**	**Decision description**
Novelty	Pacing	Data management is an established practice which is still gradually evolving and is thus a pacing field of study, but there are also emerging and disruptive components to data management
Innovation	New	The HCDMMM is a new model that draws on insights from existing maturity models in healthcare data management

5.3 Definition of the research product scope

The scope of the HCDMMM is defined with the assistance of the design decisions of Mettler (2010*b*). These design decisions are presented in Table 5.2. The scoping decisions included determining the breadth, depth and intended audience of the HCDMMM that was to be developed. The scope was further defined by determining that the human contributions, and finances and cost, that were identified in Chapter 3, are excluded from the HCDMMM development. The exclusion of human contributions, and finances and costs are discussed further in section 5.4.1.

As proposed by De Bruin *et al.* (2005) and Mettler (2010*b*), the first significant scoping decisions that set the outer boundaries for the HCDMMM application was the focus of the HCDMMM. It was decided that the focus of the HCDMMM is domain-specific (as opposed to general issue). The specific domain that the maturity model is developed for is the healthcare data management domain of developing countries. The HCDMMM was developed with the focus of being applicable to different healthcare delivery entities such as different hospitals and clinics, as well as the headquarters of the healthcare delivery organisations. In accordance with Venable and Baskerville (2012) who defined Design Science Research as research that invents new purposeful artefacts that address generalised types of problems, the HCDMMM was developed with sufficient generality to be widely applicable to data management of different healthcare delivery entities across different developing countries. Therefore, the HCDMMM components are generic in nature so as to be applicable to a variety of different healthcare delivery entities such as hospitals and clinics and their governing bodies of different developing countries.

Other design decisions stated by Mettler (2010*b*) are the depth or operating altitude of the maturity model and the particular audience the maturity model focuses on. The depth of the maturity model was decided to focus on the organisation (as opposed to individual/group, interorganisational or global/society focus). It was also decided that the maturity model will strive to represent healthcare data management on multiple system levels, so the maturity of group functions and interorganisational functions of the organisation are also addressed. Focusing on the organisational operating altitude, but also incorporating the different system levels, allowed this study to address healthcare data management on a wider spectrum and as a system.

The HCDMMM was developed according to the needs of a management-oriented audience (as opposed to technology-oriented), as managers have the knowledge to make a data management maturity assessment of their entities and can make improvement decisions based on the assessment. Developing the maturity model for a management-oriented audience means that it was developed to be understood on a strategic level. As discussed in Section 5.4.1, the HCDMMM focuses on the technical component of healthcare data management. As the audience is management-oriented, it considers the technical

Table 5.2: Scoping phase decisions (Mettler, 2010*b*)

Design parameter	**Chosen characteristic**	**Decision description**
Breadth	Specific issue	The breadth of the HCDMMM focuses specifically on data management in the healthcare of developing countries
Depth	Organisation	The depth of the HCDMMM focuses on data management in the organisation at its different facilities and its ability to cooperate with other organisations
Audience	Management oriented	Management-oriented, because management is capable of making a data management maturity assessment and making strategic data management decisions

component of healthcare data management from a strategic perspective. It describes the technical aspects of healthcare data management to be understood on a strategic level.

Descriptive purpose of use maturity models are applied to assess the as-is state of capability areas and are used as a diagnostic tool (Pöppelbuß and Röglinger, 2011). It was therefore decided that the maturity model is developed for a descriptive purpose of use and the basic design principles and descriptive purpose of use design principles described in Section 4.3.4, were considered during the development of the HCDMMM.

5.4 HCDMMM development

According to Mettler (2010*b*) this is the phase where the maturity model is developed. Following the sequence of activities for developing a maturity model as set out by De Bruin *et al.* (2005) and Becker *et al.* (2009) (see Section 5.4.2), a grid or architecture of the domain components and capability areas are first defined, and then the HCDMMM is populated. The design decisions defined by Mettler (2010*b*) were used to guide the development. Given that a top-down approach was used, the maturity levels were first defined, and then capability statements were derived to fit the definitions. After the architecture was defined with its domain components, capability areas and maturity level definitions, the HCDMMM was populated. This was an iterative design process as stated by Becker *et al.* (2009). All the different components were designed iteratively. As the HCDMMM was populated, components of the architecture of the HCDMMM were iteratively refined when it was found necessary.

Consultations were held with various SMEs in different knowledge areas to verify the design of the maturity model. After consultations with SMEs, it was also necessary to refine the HCDMMM architecture. Refinements to the architecture resulted in the need to populate the new additions, for instance the population of additional capability areas. The verification process of the HCDMMM and its different components is discussed in Section 6.2, but it is also referred to in this chapter as verification was part of the development process. The rest of this section describes the design decisions for the design phase (Section 5.4.1) and the iterative development process of the maturity model (Section 5.4.2).

5.4.1 Design decisions of the design phase

As stated by Mettler (2010*b*), the following design decisions were made for the HCDMMM design phase: (i) maturity concept; (ii) goal function; (iii) design process; (iv) design product; (v) application method and (vi) respondents. These can be seen in Table 5.3.

The maturity concept used for this maturity model is process-focused (as opposed to object or people-focused). This enabled the assessment of maturity along the process-focused dimension and therefore, the maturity levels and definitions of Paulk *et al.* (1993) could be employed. This means the HCDMMM focused on defining more effective activities and work practices. The goal function of maturity is one-dimensional (as opposed to multi-dimensional). By deciding on an one-dimensional goal function, it focuses the maturation of each capability area on a single target measure in the way that it progresses in maturity. The target measure of maturity for this study is process-focused, meaning the extent to which a specific activity is explicitly defined, managed, measured, controlled, and effective. Although the goal function for each capability area is one-dimensional, there are three maturation paths specified for the three different types of capability areas. This is discussed in Section 5.4.2.2.

Although human factors, finances and costs were found to be two significant challenge areas of healthcare data management identified in Chapter 2, they are components that fall outside of the scope of the HCDMMM, because this study focused on the continuous process improvement of the data management system. Although human factors comprise an important component of the overall system, they do not fall under continuous process improvement. Cognisance was taken of the importance of human contributions, finances and cost when developing the HCDMMM, as they were important challenges identified previously, but the HCDMMM focuses on assessing the technical component of healthcare data management. As stated in Section 5.3, the audience was decided to be management-oriented, thus the focus of this maturity model is on the technical components of healthcare data management from a managerial/strategic perspective. The HCDMMM focuses on the tech-

Table 5.3: Design phase decisions (Mettler, 2010*b*)

Design parameter	**Chosen characteristic**	**Decision description**
Maturity concept	Process	Maturity focuses on defining more effective activities and work practices. The maturity of capability areas is measured to the extent to which they are explicitly defined, managed, measured, controlled and effective
Goal function	One-dimensional	The goal function of every capability area is one-dimensional, focusing on one goal, but three maturation paths are specified for the three different types of capability areas
Design process	Combination of theory-driven and practitioner-based	Use of a structured literature review to determine domain components and use practitioners (SMEs) to refine the HCDMMM iteratively
Design product	Combination of textual description of form and functioning and software instantiation	The maturity model is presented as a textual description of form and functioning in a spreadsheet that can be opened in any software implementing the Office Open XML standard
Application method	Self-assessment	Managers of healthcare entities will use the maturity model to assess data management maturity themselves
Respondents	Combination of staff and business partners	Management of healthcare organisations like hospitals and clinics corresponds with their staff and business partners to make maturity assessments

nical component of the healthcare system, but does not go into all the micro details of the technical aspects such as specifying which software, technologies or networks to incorporate and how these elements interact with each other to be operational. The measurement of maturity of human contributions is along a different maturity dimension from what was decided on for this study, which further strengthened the reason to exclude human contributions from the developed maturity model.

With regard to the design process, it was decided to incorporate a combination of theory-driven and practice-based methods. Domain components can be identified through the use of literature (De Bruin *et al.*, 2005), but for the identification of capability areas, other qualitative methods are needed (Becker *et al.*, 2009; Lahrmann *et al.*, 2011). Therefore, interviews with SMEs were used to assist with the development of components such as the capability areas and the descriptions of maturity levels after the maturity levels were derived from literature. SMEs that were consulted had knowledge in various areas such as Structured Query Language (SQL), enterprise architecture, maturity models, and data engineering and management.

It was decided that the design product should be a combination of a textual description of form and functioning, and a software instantiation (as opposed to only a textual description of form). Making use of a textual description of form and functioning incorporates textual descriptions to describe the functioning of the HCDMMM and makes the HCDMMM easily understandable. Furthermore, it was decided that the maturity model was to be presented in a spreadsheet that can be opened by Office Open XML standards compliant open-source software like Libre Office or proprietary software like MS Excel. This enables the simple assessment of capability areas through its formula functionality, and as it is widely used and understood in industry.

The method of application was decided to be self-assessment (as opposed to third party assisted or done by certified professionals), as the management of the different healthcare entities would use the HCDMMM themselves to make data management maturity assessments. This would enable management to make data management assessments without needing the aid of third parties. This decision would also affect the design of the HCDMMM as it would include all the necessary information that allows management to use the HCDMMM without needing prior training.

With regard to the respondents that should be included by management to make maturity assessments, the management can liaise with their staff and business partners to determine the level of data management maturity. By doing this, management can make even more insightful assessments of their data management maturity.

5.4.2 Iterative development process of the various components

An iterative design and populate process was followed to develop the various generic components of the healthcare data management maturity model so that it is applicable to different healthcare entities such as different hospitals and clinics and healthcare delivery organisations' headquarters. For each iteration, the architecture of the HCDMMM was first determined or refined. This included domain components, capability areas, and maturity levels. The architecture was then populated with maturity level descriptions for each of the capability areas of the domain components. After the HCDMMM was developed for an iteration, its components were verified to determine whether they were theoretically sound or whether another iteration was required. The developed model was verified through the assistance of literature and SMEs. An updated architecture of the HCDMMM was then constructed based on the contributions from literature and/or SMEs for another iteration of the design process. After the architecture was refined for another iteration, the HCDMMM was repopulated to derive a model that was an improvement on the previous iteration. Each iteration built on the preceding iteration to refine the domain components, capability areas, system levels and maturity levels using the insights of the different SMEs as they were introduced to the development process. A summary of the changes that occurred during every iteration can be seen in Appendix D.

The rest of this section describes the development of the different components of the maturity model. This includes the development of the domain components (Section 5.4.2.1), maturity levels (Section 5.4.2.2), system levels (Section 5.4.2.3) and capability areas (Section 5.4.2.4).

5.4.2.1 Domain components

Due to the fact that the healthcare data management domain is complex, a stage-gate approach is used to develop it, as this approach allows for further detailed assessment of the specific domain components of healthcare data management. The different layers of the HCDMMM are the domain, domain components and domain sub-components, which are also called capability areas. This enables separate maturity assessments for the discrete domain components. The collective of these components gives an overall assessment of the entity under study.

The domain focus is outlined in Section 5.3. As De Bruin *et al.* (2005) and Becker *et al.* (2009) proposed, the domain components were derived from literature following an iterative process. Literature on big data value chains was reviewed in Section 3.1.1 and the primary value chain components from the big data value chain were used to determine the maturity model domain components. The healthcare data value chain in Section 3.1.1 depicts eight primary value chain components. Value chain components do not necessarily happen independently of each other and they often overlap. De Bruin *et al.* (2005) advised keeping the number of domain components and sub-components low to minimise perceived complexity in the HCDMMM. For this reason, the primary value chain components that overlapped were combined to reduce the number of domain components.

The data collection domain component relates to the primary value chain component data collection. The data storage domain component relates to the data sources value chain component. The data sharing domain component consists of the data acquisition and data transmission value chain components and the data analysis domain consists of the data storage for analysis and data analysis. The data usage domain component relates to the data usage value chain component. Data curation maintains the quality of data at different domain components such as at collection, storage, before sharing and before analysis. Therefore, data curation was not included as a separate domain component, but was included under these domain components. The secondary value chain components, technology and infrastructure, were included as capability areas under all the different domain components as discussed in Section 5.4.2.4. Finances and costs, and human factors were excluded from the scope of the HCDMMM as discussed in Section 5.4.1.

The value chain in Chapter 3 also depicted data governance as a key component of data management, which is not part of the primary value chain, but part of the enabling environment. To ensure that the HCDMMM is collectively comprehensive, data governance was also included. These domain components for healthcare data management were confirmed by various SMEs as described in Chapter 6.

The domain components developed for this study through the iterative development process and with the assistance of literature and the input from SMEs are: (i) data collection; (ii) data storage; (iii) data sharing; (iv) data analysis; (v) data usage; and (vi) data governance. The development of the capability areas of these domain components is described in Section 5.4.2.4. These domain components are described in the following list:

1. Data collection: the collection of data from the real world, such as from patients on the facility level for patient care delivery, and data collection for healthcare management and decision-making on the organisational level.

2. Data storage: patient data and other data collected at the facility is stored at the facility and made available through data sharing for use in various ways. Collected data for use on the organisational level is also stored for management and decision-making.

3. Data sharing: data stored at the facility is available at other facilities so that when patients visit other facilities, continuous care can be provided

without collecting data again, which can cause data duplication. Data is shared for analysis and other purposes. The eventual goal is that patients' data is available to authorised users organisation-wide.

4. Data analysis: data analysis is carried out on the facility level and on the organisation level. The facility analyses data for decision-making, operations and patient care on the facility level. Data analysis on the organisational level uses data from different facilities for decision-making that affects the whole organisation and its management.

5. Data usage: data is used in various different ways. Collected, stored, shared and analysed data is used for various different purposes to support healthcare delivery and decision-making.

6. Data governance: these domain components determine the strategy of healthcare data and develop regulations, policies and standards that direct the management of healthcare data.

5.4.2.2 Maturity levels

In Section 5.4.1 it was decided that the maturity concept should be process-focused and that the goal function is one-dimensional. The goal function focuses on the improvement of the efficiency of the processes and activities specified, but different types of capability areas have different maturation paths. These different types of capability areas are described in Section 5.4.2.4.

A logical progression through the maturity stages is of key importance (De Bruin *et al.*, 2005; Röglinger *et al.*, 2012). These stages were derived iteratively. As domain components evolved or when it was found that the maturation path was not logical, or the incremental steps were not adequately distinct, maturity levels were refined. This ensured that continuous process improvement is based on small, evolutionary steps. The iterative determination of maturity levels also ensured distinct and well-defined maturity stages.

The maturity levels stated by Paulk *et al.* (1993) are used to determine the maturity levels for this study as they provide a logical and anticipated progressive maturation through maturity stages for processes and activities. As the HCDMMM strives to address a generalised type of problem (Venable and Baskerville, 2012), whether a process or activity is carried out on paper or electronically was not specified in the maturity levels. The HCDMMM was developed to assess maturity according to how well the identified capabilities are performed according to the Capability Maturity Model (CMM) levels of Paulk *et al.* (1993) as described in Section 4.2.3, agnostic of whether it is carried out on paper or electronically. Specifying that a capability area is on a certain maturity level depending on whether it is paper-based or electronic, implies that changing from a paper-based process to an electronic process guarantees improvement in how well the function or process is performed, which is not in reality the case. A system can still be paper-based and perform a function better than one done electronically. This is not always the case, but it was found that maturity is not necessarily dependent on whether the process or activity is carried out on paper or electronically. Therefore, to maintain the generalised quality of the HCDMMM, the maturity levels were developed generically.

The resultant maturity levels that were implemented across the different domain components and system levels were classified as *Initial*, *Repeatable*, *Defined*, *Managed* and *Optimising*. These maturity levels describe how well the different specified capability areas are performed within the different domain components.

Three maturation paths were differentiated for the different capability area types according to the three data value chain levels defined in Chapter 3. These levels are the primary value chain, the secondary value chain and enabling environment components. The primary value chain components consist of primary activities. The secondary value chain consists of supportive components to the primary activities and the enabling environment consists of the enabling practices. Because the components of these three value chain categories differ, the logical and anticipated way that they progress towards maturity will also differ. Thus, three main maturation paths were determined for the different components according to the value chain level categories. A maturation path was determined for the primary activities of each domain component, another one for the supportive structures of the primary activities and one that indicates the maturation of enabling practices. All capability areas of the different domain components fall under one of these three capability categories with their different maturation paths. The maturity level definitions of the three differentiated maturation paths for primary activities, supporting structures and enabling practices are defined in Table 5.4, Table 5.5 and Table 5.6, respectively.

Table 5.4: Maturity level definitions for primary activities (adapted from Paulk *et al.* (1993))

Level	Level name	Level definition
1	Initial	Primary activities are not stable and controlled. Objectives are reached late and at high cost. Execution of primary activities is dependent on individuals. Effective execution is sporadic and not the norm
2	Repeatable	Primary activities are carried out in a disciplined way and can be carried out repeatedly. The primary activities are stable and repeatable

Continued on next page

Continued from previous page

Level	**Level name**	**Level definition**
3	Defined	Primary activities are documented. Primary activities are integrated into a coherent whole. Primary activities are established for effectiveness. Primary activities are well-defined with inputs, standards and procedures for performing functions and outputs. Activities are standard, consistent and quality is checked
4	Managed	Primary activities are measured/ monitored/ controlled/ supervised/ organised in accordance with the defined practices through available data for measurement. Primary activities are quantifiable and predictable. Primary activities are stable and are measures that make it possible to address special cases and take corrective action
5	Optimising	Primary activities are continuously improved. Weaknesses of primary activities are identified and proactively strengthened. The method and range of the ability of the primary activity is improved. Improvements are planned and managed and are a part of normal business activities. Innovations that exploit best practices are identified and transferred after cost-benefit analysis is done on data of primary activities

Table 5.5: Maturity level definitions for supporting structures (adapted from Paulk *et al.* (1993))

Level	**Level name**	**Level definition**
1	Initial	Initial supportive structures for data management exist. Initial structures give very limited support to the functioning of primary activities
2	Repeatable	Appropriate supportive structures that give adequate support to data management for repeatable primary activities are applied
3	Defined	Specific supportive structures that give effective support to data management primary activities are applied. Support structures support the functioning of primary activities as a coherent whole
4	Managed	The availability, reliability and maintainability of supportive data management structures are monitored. Support structures are repaired or replaced when failure occurs

Continued on next page

Continued from previous page

Level	**Level name**	**Level definition**
5	Optimising	Supportive structures for data management are continuously maintained and upgraded whenever needed. Support structures are proactively maintained or replaced

Table 5.6: Maturity level definitions for enabling practices (adapted from Paulk *et al.* (1993))

Level	**Level name**	**Level definition**
1	Initial	Enabling practices do not create a stable environment to develop and maintain primary activities and support structures. Enabling practices do not control primary activities and support structures
2	Repeatable	Basic enabling practices for primary activities and supporting structures are established and procedures to implement policies are available. Enabling practices create a stable environment for primary activities and supporting structures
3	Defined	Well-defined enabling practices like standards exist. Organisation-wide understanding of the activities, roles and responsibilities are in place
4	Managed	Enabling practices exist for quantifying primary activity goals and measurements. Enabling practices exist to measure/ monitor/ control/ supervise/ organise support structures. Established enabling practices are measured/ monitored/ controlled/ supervised/ organised
5	Optimising	Enabling practices of data management are continuously upgraded and updated. Existing enabling practices enable the continuous improvement of primary activities and the implementation of the support structures to support the improvements

5.4.2.3 System levels

Given that the HCDMMM focuses on the healthcare data management system, it is necessary to integrate systems approach components that were stipulated in Section 2.1.1 into the design of the maturity model to represent the healthcare data management system comprehensively. As discussed in Section 2.1.2, Ferlie and Shortell (2001) proposed four levels in the healthcare system to consider when successful change towards better quality care is desired. These levels are: (i) the individual patient; (ii) the care team; (iii) the overall organisation; and (iv) the environment in which the organisation is embedded. These levels are nested and each has an influence on the other levels. Different components on the levels interact with each other to operate as sub-systems in the system of interest. Different capability areas of the healthcare data management system are carried out on different system levels.

This model focuses on two of these system levels, which are the care team (or facility level) and the organisational level, which were developed while considering the influence of the individual patient and the environment on these levels. The facility level is related to the care team level, because care is delivered from healthcare facilities. The facility is the microsystem and the replicable unit within the organisation that contains within itself all the necessary resources to do its work. This relates to entities such as hospitals and clinics. The facility level is nested in the organisational level. The organisation coordinates the activities of multiple facility care units through: (i) decision-making systems; (ii) information systems; (iii) operating systems; and (iv) processes like financial, administrative and human resources. The economic and political environment influences the organisation directly and influences all the other levels indirectly through its influence on the organisation. The HCDMMM was designed to consider the influences of the environment on the organisation. Given that data management supports the delivery of healthcare to the patient, the capability areas of the facility level considered the needs of the patient when the components of the facility level were developed.

It is also noteworthy that primary activities, supporting structures and enabling practices, as described in Section 5.4.2.4, are different capability areas that work together to perform certain functions in the system. These capability categories relate to different system components that interact with each other on various levels to perform certain outcomes. Enabling practices has a regulating influence on primary activities and support structures, and the support structures are the necessary components that need to be in place in order to carry out the primary activities.

5.4.2.4 Capability areas

It is not very likely that literature will be sufficient to determine domain capability areas and exploratory research methods are recommended to determine them (De Bruin *et al.*, 2005; Becker *et al.*, 2009; Lahrmann *et al.*, 2011). The capability areas were derived from knowledge gained from literature on the respective domain components and were populated iteratively. Capability areas for the different domain components were derived from literature and verified by SMEs after different development iterations. SMEs were consulted to identify relevant capability areas that were not identified through literature. This ensured that the included capability areas were verified and that important capability areas were not omitted.

A single maturation path does not adequately describe the maturity progression of all the capability areas as they consisted of different healthcare data value chain levels and it was necessary to distinguish different capability categories to base different maturation paths on. Three differentiated capability categories were specified. These capability categories were primary activities, supporting structures and enabling practices which were also explained under Section 5.4.2.2. Table 5.7 illustrates the different capability categories at the different domain components and system levels.

These capability categories are distinguishable through their focus of purpose. The primary activities are the core activities in every domain component to accomplish the goal of that specific domain component. Supporting structures are the secondary value chain components that support the working of the core activities to accomplish the domain component goal. These capability areas are the specific secondary inputs and services needed to accomplish the goals of the primary activities. The enabling practices are the conventions that shape the working of the primary activities and supporting structures. These are the customs that create the environment that the primary activities and supporting structures operate in. Table 5.7 indicates how the three different capability areas relate to the different domain components and system levels.

To determine the different capability areas of the domain components, the design protocol as outlined in this section was followed. The design protocol was also carried out iteratively for each domain component to ensure all the necessary capability areas for that domain component were included. The design protocol was used for every domain component, except for the data governance domain component, which only consists of enabling practices. The design protocol was used to determine the data collection, the data storage, the data sharing, and the data usage capability areas. To derive data governance capability areas, only steps three and five of the design protocol were carried out. For each iteration of the HCDMMM development, the design protocol to determine the capability areas of each domain component was used. Not all the steps of the design protocol were necessary to be carried out during every iteration of the HCDMMM development and were skipped accordingly. The steps of the design protocol are:

Table 5.7: The different system levels and capability categories per domain component

Domain component	System level	Capability category
Data collection	Organisation	Primary activities
		Supporting structures
		Enabling practices
	Facility	Primary activities
		Supporting structures
		Enabling practices
Data storage	Organisation	Primary activities
		Supporting structures
		Enabling practices
	Facility	Primary activities
		Supporting structures
		Enabling practices
Data sharing	Organisation	Primary activities
		Supporting structures
		Enabling practices
	Facility	Primary activities
		Supporting structures
		Enabling practices
Data analysis	Organisation	Primary activities
		Supporting structures
		Enabling practices
	Facility	Primary activities
		Supporting structures
		Enabling practices
Data usage	Organisation	Primary activities
		Supporting structures
		Enabling practices
	Facility	Primary activities
		Supporting structures
		Enabling practices
Data governance	Organisation	Enabling practices

1. Derive the main primary activities from the knowledge gained from literature for the specific domain component that accomplish the purpose of the domain component. This includes:

   1.1. determining on what system level(s) they are carried out;

   1.2. adding a capability row for each activity at the appropriate system level(s); and

1.3. populating all the different maturity levels of the different primary activities identified with descriptions according to the defined maturity level definitions for primary activities.

2. Identify the supporting structures for the specific domain component that supports the primary activities to accomplish the purpose of the domain component. This includes:

    2.1. determining on what system level(s) they are implemented to support the primary activities;
    2.2. adding a capability row for each supporting structure at the appropriate system level(s); and
    2.3. populating all the different maturity levels of the different supporting structures identified with descriptions according to the defined maturity level definitions for supporting structures.

3. Identify the enabling practices of the domain component that determine the conventions that shape the working of primary activities and supporting structures. This includes:

    3.1. determining on what system level(s) they are implemented to enable the primary activities and supporting structures;
    3.2. adding a capability row for each enabling practice at the appropriate system level(s); and
    3.3. populating all the different maturity levels of the different enabling practices identified with descriptions according to the defined maturity level definitions for enabling practices.

4. Add a generic capability row for the standards adherence of all domain component activities at the system level(s) where it is applicable. Populate the capability row so that it describes the maturity of the adherence to the standards for all the primary activities and supporting structures in general terms.

5. Verify, and revise where necessary, the HCDMMM components and maturity descriptions with the support of literature and SME contributions.

The maturity model describes primary activities that are specific to each domain component. The primary activities of each domain component are distinct. These primary activities ensure that the overall goal of the domain component is achieved.

Many enabling practices (i.e. data entry forms/structure, types of data captured, retention plan of stored data etc.) and supporting structures (i.e. data transmission networks, data queries for various usages, etc.) are also specific to a domain component and are distinct; however, some are recurrent across the different domain components. These capability areas are stated under each applicable domain component. Because the focus of their function is specifically different for each domain component, they need to be assessed under each domain component to obtain an accurate maturity score for each of them. These recurrent capability areas are ones that focus on technology and infrastructure, data privacy and security, and alignment to standards, policies and regulations. All the capability areas with their maturity level descriptions can be viewed in Appendix E.

5.5 The HCDMMM

In this section the developed HCDMMM is presented. In Section 5.5.1 the conceptual HCDMMM is presented and in Section 5.5.2 the HCDMMM is presented in the format of how it is transferred to users.

5.5.1 The conceptual HCDMMM

The design and development methodology as described in Section 5.4 resulted in the Healthcare Data Management Maturity Model (HCDMMM) with the components as described in Sections 5.4.2.1, 5.4.2.2, 5.4.2.3 and 5.4.2.4. It can be used to determine the maturity scores of the overall data management system of a healthcare entity within developing countries and does not assess individual data management systems within the entity. It consists of four components to assess the as-is state of healthcare entities' overall data management related to patient care delivery. These components include: (i) two system levels (i.e. the facility and organisational levels); (ii) six domain components (i.e. data collection, data storage, data sharing, data analysis, data usage and data governance); (iii) three capability categories that contain all the capability areas (i.e. primary activities, supporting structures and enabling practices); and (iv) five maturity levels (i.e. *Initial*, *Repeatable*, *Defined*, *Managed* and *Optimising*). A healthcare entity's data management can either be assessed on the facility or organisational level (as determined in Section 5.4.2.3), depending on which level of the system it operates on. The HCDMMM can be used to evaluate the data management of the hospitals and clinics on the facility level and to evaluate the data management of the headquarters of a healthcare organisation. Figure 5.2 illustrates the conceptual maturity model and how all its different components fit together.

The domain components are the overarching maturity assessment categories that consist of multiple capability areas that are assessed individually to determine the overall maturity of the domain component. The six domain components for this study were determined in Section 5.4.2.1. The five maturity levels are used to assess the maturity of each capability area. Each

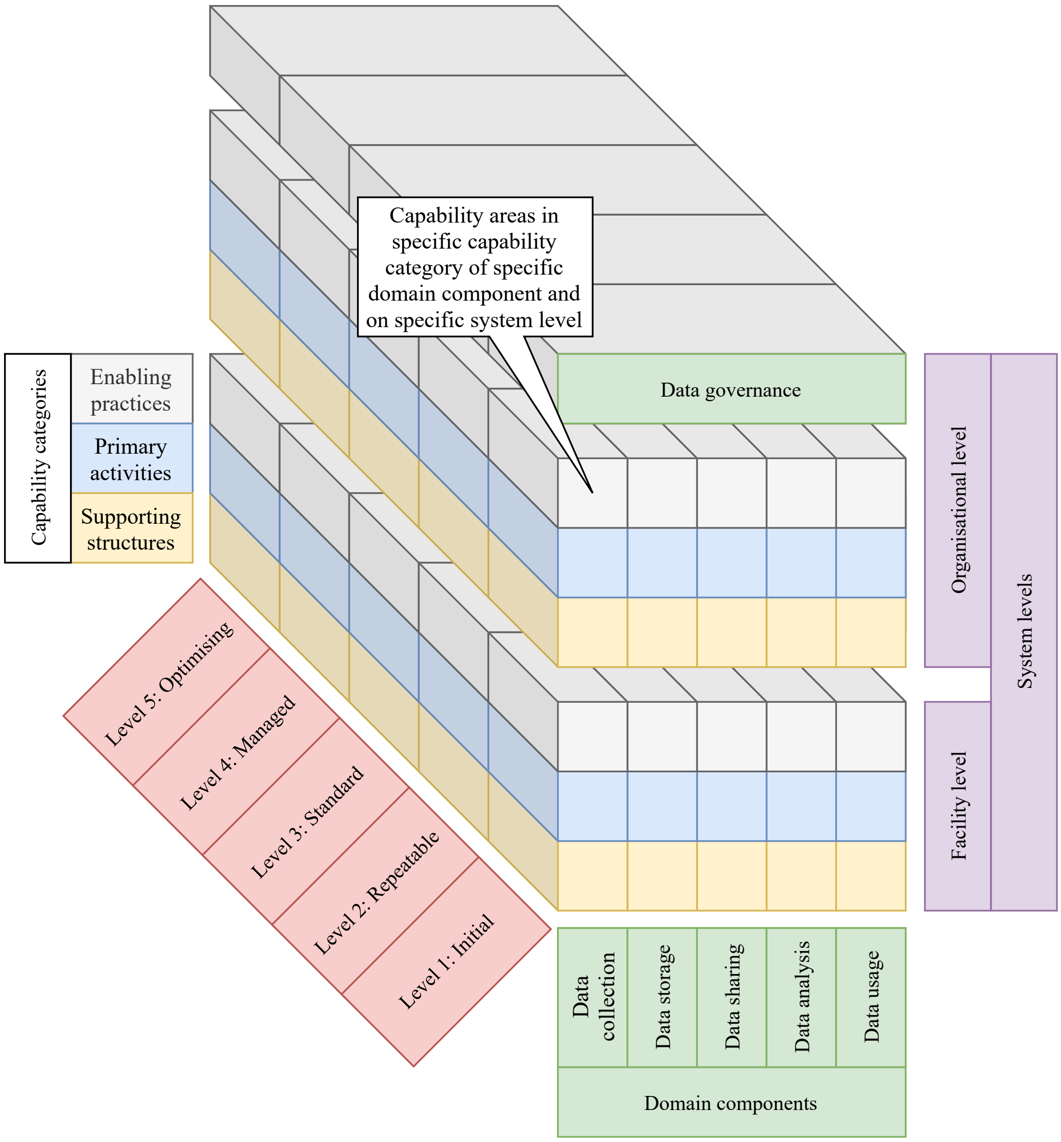


Figure 5.2: Conceptual maturity model (adapted from Van Dyk (2013))

capability area is described by its own set of maturity level descriptions. Capability areas belong to one of three different capability categories that determine their maturation paths. This is due to different types of capability areas maturing differently. These three different maturation paths were determined in Section 5.4.2.2. Each of the three maturation paths has five maturity levels and they were used to determine the maturity descriptions of all the capability areas. The maturity level descriptions of all the capability areas can be seen in Appendix E.

The intended users of the HCDMMM are individuals in management positions that have sufficient knowledge and understanding of their healthcare entity's data management in order to be able to use the tool to assess it, whether the entity is on the facility or organisational level. The users that do the assessments can either be individuals that have sufficient knowledge of the data management system, or a group of individuals whose collective knowledge covers the data management system of the entity. If management does not have adequate knowledge of the entity's data management to make an assessment, they can consult IT personnel and other healthcare workers (HCWs) that work with the data management system in order to be able to make an assessment.

Comparing the actual functioning of components of the real-world system with the maturity level descriptions of the capability areas allows the maturity assessment of the processes, objects or functions of the real-world system with regard to each capability area. The HCDMMM is not used to assess individual services of the entity under study, but to give an overall assessment of every capability area in the different domain components.

After the components of the HCDMMM were defined through the iterative design process based on theoretical and SME inputs, the next stage was the conception of the transfer media and evaluation (Becker *et al.*, 2009; Pöppelbuß and Röglinger, 2011). During that stage, the form in which the maturity model will be transferred to user communities is determined and developed, as well as how users should use the maturity model to make a maturity assessment. This involved the transformation of the developed components through the iterative design process into a functioning and usable maturity assessment tool. The HCDMMM is presented in the following section in the format of how it is transferred to users, after both the verification and validation processes were conducted as described in Chapter 6.

5.5.2 Conception of transfer media of the HCDMMM

As stated previously, the transfer media is the form in which the maturity model will be transferred to user communities. It also instructs users on how they should use the maturity model to make a maturity assessment. The transfer media of the HCDMMM was developed in a macro-enabled workbook that consists of 18 sheets. The transfer media workbook can be opened in any software that implements the Office Open XML standard. The assessment tool was designed to assist healthcare entities such as hospitals, clinics and headquarters to assess their current state of data management in order to determine which components can be improved on and how. The tool is meant to be used in two scenarios: (i) before improvement endeavours to identify components of improvement; and (ii) after improvement endeavours to determine whether the implementations for improvement actually realised their goal. The tool was designed in such a way as to be easily usable to the intended user, with simple navigation between assessment sheets that can be revisited when needed. The tool also seeks to provide the assessment results in an easy-to-understand format, with visual representations of the results. The HCDMMM is described with the aid of Figures 5.3 to 5.12. The results sheet with hypothetical results are illustrated in Appendix F.

The landing page consists of an introduction to the HCDMMM, which can be seen in Figure 5.3. The purpose of the introduction is to give the user background information on why the tool was developed, the aim of the HCDMMM, the general purpose of maturity models and how the tool was developed. The introduction also includes which types of entities can be assessed by using the tool and it also specifies who the intended users of the HCDMMM are.

The landing page is followed by an overview of the HCDMMM. This includes the level of analysis of the HCDMMM, how it is structured and what components it includes. The content of this overview sheet can be seen in Figure 5.4. The overview also describes some maturity model concepts so that the users have a better understanding of maturity models before doing the assessment. The HCDMMM strives to use standard domain language so that users can easily understand what is meant by the HCDMMM, but many terms can be ambiguous and it was therefore necessary to establish a common understanding of maturity model concepts used in the HCDMMM. Terminology used to describe different maturity levels was also defined so that all users have a common understanding of those terms before assessments are done as can be seen in Figure 5.6. The concept of maturity is described and the maturation paths of the HCDMMM are defined. The three defined maturation paths give the user an understanding of how capability areas of different capability categories progress. The maturation paths are illustrated in Figure 5.5.

Introduction

HCDMMM aim

The **Health Care Data Management Maturity Model** (HCDMMM) is a maturity assessment tool that was developed to assist the assessment of the overall data management of healthcare entities on a facility level such as hospitals and clinics, as well as on an organisational level such as the headquarters of a healthcare delivery organisation. Data management assessments are done per healthcare entity. The HCDMMM includes data management components related to patient care delivery on both these levels. Data management in developed countries strive to address the problem of preventative care, where most developing countries struggle with data management that supports the treatment of disease. The healthcare data management activities of the HCDMMM are more suitable for developing countries and thus, the HCDMMM focuses on data management of developing countries.

Importance of data management in healthcare

Data management assists the effective care delivery to patients and helps management with decision-making on numerous levels. If a healthcare entity's data management practices are insufficient, it can inhibit the entity's ability to provide effective care to patients and wrong data can cause decision-makers to make decisions based on incorrect data. Therefore, it is very important to improve healthcare entities' data management to ensure that they support care delivery to patients. The HCDMMM assesses an entity's data management that can be used as a foundation to improve the entity's data management as it determines on what level of maturity the entity's data management is, and identifies the weakest data management components of the entity as potential focus points for improvement.

Overview of maturity models

A maturity model is designed to assess the maturity of a selected domain based on a set of criteria. They are artefacts that determine the *status quo* of the capabilities of an entity. A maturity model consists of stages for each capability area that forms an anticipated, desired or logical path from an initial to a target maturity state. By using this maturity assessment tool, an healthcare entity can:

- Identify its current overall data management maturity level
- Identify the maturity levels of the main data management components
- Identify specific areas that are the weakest and that are potential focus points for improvement

The development process of the HCDMMM

The HCDMMM was developed by incorporating knowledge gained from a multidisciplinary literature review that covers components of data management, healthcare systems and maturity models. This served as a solid foundation of the HCDMMM. The HCDMMM was further refined using inputs of subject matter experts in these various knowledge domains.

The model was developed with enough generality to be widely applicable to the data management of different healthcare entities, but also with sufficient specificity that the model is of beneficial use and that it can assess specific components of the entity's data management. The model focuses on the assessment of data management from a strategic perspective, which is another reason why it was not developed with greater specificity, especially regarding technological components.

The healthcare delivery entities that the HCDMMM focuses on

The model can be used to assess two types of entities on two levels of a healthcare delivery system; the facility and the organisational levels. The HCDMMM focuses on data management activities related to care delivery to patients. These data management practices differ for entities on the facility and organisational level. Therefore, before an entity can assess its data management, it should specify on which system level the assessment will take place.

The facility level is nested in the organisational level. Therefore, the entity on the organisational level has an influence on the entities that are nested in it. The facility level is related to the care team level, because care is delivered from healthcare facilities. The facility is the microsystem and the replicable unit within the healthcare delivery system that contains within itself all the necessary resources to do its work. Therefore, facility entities such as hospitals and clinics can use this model.

The entity on the organisational level is responsible for the entities on the facility level and has a regulatory and managerial influence on its facilities. Therefore, entities such as the headquarters of healthcare delivery organisations can use the HCDMMM to make maturity assessments of their data management.

The economic and political environment directly influences the entity on the organisational level and all other facilities indirectly through its influence on the entity on the organisational level. The model was designed to consider the influences of the environment on the organisational level.

The intended users

The intended users of the HCDMMM are individuals in management positions that have sufficient knowledge and understanding of its healthcare delivery entity's data management in order to be able to use the tool to assess its data management, whether the entity is on the facility or organisational level. The users that do the assessments can either be an individual that has sufficient knowledge of the whole data management system, or a group of individuals whose collective knowledge covers the whole data management system of that entity. If management does not have adequate knowledge of the entity's data management to make an assessment, they can consult IT personnel and other healthcare workers that work with the data management system in order to be able to make an assessment.

The assessment of healthcare data management systems can also be assisted by third-parties such as the World Health Organization or the Global Fund who are invested in the improvement of healthcare in developing countries. These organisations are also interested in the assessment results as they will be able to find trends and patterns across facilities and can identify which capability areas should be focused on in general for investment and improvement endeavours.

Model overview

Figure 5.3: Landing page of the HCDMMM

Model overview

Level of analysis
The HCDMMM is structured to be understood and assessed on a strategic level. The use of this model gives the user an overall assessment on a strategic level of capability areas in healthcare data management. It does not result in a detailed assessment of capability areas for specific data management systems within the entity.
As stated previously, the HCDMMM can be used to assess entities on the facility or the organisational level. The organisational level is the macro level of the healthcare data management system where the macro-functions are carried out. This is the managerial level and functions on this level often affect functions on the micro level too. The functions on this level have an organisation-wide effect. The facility level is the micro and operational level of the healthcare data management system where micro functions are carried out. The facility is the micro system and the replicable unit within the healthcare delivery system that contains within itself all the necessary resources to do its work. The HCDMMM presents each of these two levels via their own set of Excel worksheets.

Domain components
There are six domain components of healthcare data management that the HCDMMM assesses and each domain component is presented on its own sheet.
Domain components are significant particular parts of a specific domain that play a significant role in the maturity of the domain. These are high-level categories for the areas that are to be assessed. The six domain components included in the model pertaining to healthcare data management are:
- Data collection: the collection of data from the real world, such as from patients on the facility level for patient care delivery, and data collection for healthcare management and decision-making on the organisational level.
- Data storage: patient data and other data collected at the facility is stored at the facility and made available through data sharing for the usage of data in various ways. Collected data for use on the organisational level is also stored for management and decision-making.
- Data sharing: data stored at the facility is available at other facilities so that when the patient visits other facilities, continuous care can be provided without collecting data again, causing data duplication. Data is shared for analysis and other data purposes. The eventual goal is that data of patients are available organisation-wide.
- Data analysis: data analysis is carried out on the facility level and on the organisation level. The facility analyses data for decision-making, operations and patient care on the facility level. Data analysis on the organisational level uses data from different facilities for decision-making that affects the whole organisation and its management.
- Data usage: data is used in various different ways. Collected, stored, shared and analysed data is used for various different purposes to support healthcare delivery and decision-making.
- Data governance: the data governance domain component is the domain component that determines the strategy of healthcare data and develop regulations, policies and standards that direct the management of healthcare data.

Capability areas
Distinct capability areas exist within the domain components that allow for further detailed assessment of the domain components. A capability area relates to the ability of a system to perform a specific function, process or cluster of activities. The HCDMMM specifies three types of capability areas according to the role they fulfil in domain components. These are:
- Primary activities: The primary activities are the core activities in every domain component to accomplish the goal of that specific domain component.
- Supporting structures: Supporting structures are the secondary capability areas in the domain components that support the working of the core activities to accomplish the domain component goal. These capability areas are the specific secondary inputs and services needed to accomplish the goals of the primary activities.
- Enabling practices: The enabling practices are the conventions that shape the working of the primary activities. These are customs that create the environment that the primary activities are carried out in.
The different capability areas of the HCDMMM are presented per domain component.

Assessment results
The HCDMMM presents the assessment results after all the domain components are assessed. Results are presented per domain component, per capability category and the user can create customised views of capability scores according to certain criteria.

Instructions | Back to introduction

Figure 5.4: Model overview

Maturation paths

The capability areas are assessed along maturation paths according to their capability type. A maturation path is the anticipated, desired or logical path a capability area undergoes towards a target maturity state from an immature state. It consists of a finite number of maturity levels that describe how well the capability areas are anticipated to perform at each level. The defined maturation paths for primary activities, supporting structures and enabling practices can be seen below. These determine the maturity descriptions of the capability areas under each of these three capability categories. The HCDMMM strives to make a healthcare data management assessment agnostic of whether activities are carried out on paper or electronically.

Primary activities (PAs)					
Level	1	2	3	4	5
Name	**Initial**	**Repeatable**	**Defined**	**Managed**	**Optimising**
Definition	Primary activities are not stable and controlled. Objectives are reached late and at high cost. Execution of primary activities is dependent on individuals. Effective execution is sporadic and not the norm	Primary activities are carried out in a disciplined way and can be carried out repeatedly. The primary activities are stable and repeatable	Primary activities are documented. Primary activities are integrated into a coherent whole. Primary activities are established for effectiveness. Primary activities are well-defined with inputs, standards and procedures for performing functions and outputs. Activities are standard, consistent and quality is checked	Primary activities are measured/ monitored/ controlled/ supervised/ organised in accordance with the defined practices through available data for measurement. Primary activities are quantifiable and predictable. Primary activities are stable and are measures that make it possible to address special cases and take corrective action	Primary activities are continuously improved. Weaknesses of primary activities are identified and proactively strengthened. The method and range of the ability of the primary activity is improved. Improvements are planned and managed and are a part of normal business activities. Innovations that exploit best practices are identified and transferred after cost-benefit analysis is done on data of primary activities

Supporting structures (SSs)					
Level	1	2	3	4	5
Name	**Initial**	**Repeatable**	**Defined**	**Managed**	**Optimising**
Definition	Initial supportive structures for data management exist. Initial structures give very limited support to the functioning of primary activities	Appropriate supportive structures that give adequate support to data management for repeatable primary activities are applied	Specific supportive structures that give effective support to data management primary activities are applied. Support structures support the functioning of primary activities as a coherent whole	The availability, reliability and maintainability of supportive data management structures are monitored. Support structures are repaired or replaced when failure occurs	Supportive structures for data management are continuously maintained and upgraded whenever needed. Support structures are proactively maintained or replaced

Enabling practices (EPs)					
Level	1	2	3	4	5
Name	**Initial**	**Repeatable**	**Defined**	**Managed**	**Optimising**
Definition	Enabling practices do not create a stable environment to develop and maintain primary activities and support structures. Enabling practices do not control primary activities and support structures	Basic enabling practices for primary activities and supporting structures are established and procedures to implement policies are available. Enabling practices create a stable environment for primary activities and supporting structures	Well-defined enabling practices like standards exist. Organisation-wide understanding of the activities, roles and responsibilities are in place	Enabling practices exist for quantifying primary activity goals and measurements. Enabling practices exist to measure/ monitor/ control/ supervise/ organise support structures. Established enabling practices are measured/ monitored/ controlled/ supervised/ organised	Enabling practices of data management are continuously upgraded and updated. Existing enabling practices enable the continuous improvement of primary activities and the implementation of the support structures to support the improvements

Figure 5.5: The maturation paths on the HCDMMM overview sheet

Terminology definitions

Definitions of terminology that is used in the maturity levels are clarified before the maturity assessment is done so that all HCDMMM users have a common understanding of terminology used to describe the different maturity levels.

Terminology	Description
Adequate	Supports/enable primary activity so that it can perform its basic function
Appropriate	Suitable or proper in the circumstances
Available	At the intended user's disposal without difficulty when necessary
Baseline performance	A minimum or starting point performance
Comprehensive	Including or dealing with all, or nearly all, elements or aspects of an area
Effective	Successfully reaching the intended result
Efficient	Achieving the maximum output with minimum input
Entity	The distinct entity being assessed by the tool, whether it is an entity on the facility level like a hospital or clinic, or an entity on the organisational level like a healthcare delivery entity's headquarters
Reliable	The consistent performance of basic functions
Resource-intensive	The high number of the collective effort of humans, technology and cost
Robust	Supporting structures that are long-lasting and are able to maintain its operation without failure
Structured data	Human- or machine-generated data with clearly defined data types that resides in relational databases
Unstrucutured data	Textual- or non-textual, human- or machine-generated data with internal structure but that is not structured via pre-defined data models or schema including formats like images, audio and video

Figure 5.6: Definitions of terminology used in maturity levels

The instructions, which are depicted in Figure 5.7, can be navigated to subsequent to the HCDMMM overview. The instructions describe how the HCDMMM can be put into action to make a maturity assessment. The instructions ensure that the user understands how to use the HCDMMM without the need for training, as the instructions, along with the overview data, aim to provide sufficient information to users to use the tool. The instructions also instruct the user on how to navigate between sheets to ensure ease of navigation. Furthermore, it instructs the user on how to address uncertainty when doing the assessment. At the end of the instruction sheet, the user should specify on which system level the assessment will take place so that they are automatically directed to the appropriate set of sheets.

From there, the user can navigate to either the organisational level overview or the facility level overview. These overviews can be seen in Figure 5.8 and 5.9, respectively. As the HCDMMM can assess the data management of healthcare entities on the organisational and facility level, the sheets are split into the assessment of either of the two. This means that the sheets that are described next exist for the organisational level and the facility level. Figures 5.10 to 5.12 focus on exhibiting content for the organisational level, but the format and structure is the same for the facility level; the content is just different to accurately represent entities on the specific system level.

Once the user has specified on the instruction sheet on which system level the assessment will take place, the HCDMMM will direct them to an overview of the capability areas that are going to be assessed for that specific system

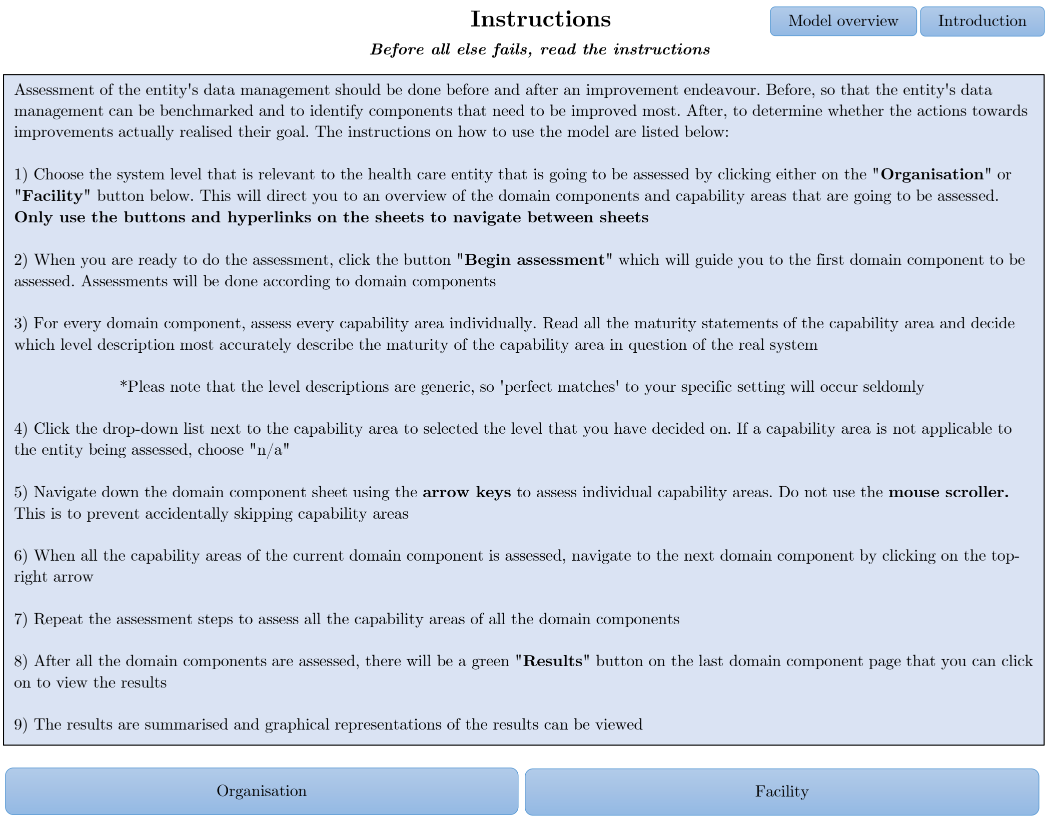


Figure 5.7: Assessment tool instructions

level. These overviews can be seen in Figure 5.8 and 5.9. These high-level overviews give the user the extent of the HCDMMM content for the system levels and provide them an indication of the knowledge that will be needed to do the assessment. The capability areas are categorised according to their domain component. There is a key that indicates which capability category each capability area belongs to. From there the user can start the assessment and the HCDMMM will navigate to the first domain component so that all its capability areas can be assessed.

The assessment is carried out per domain component. After all the capability areas of the domain component are assessed, the user can navigate to the next domain component to assess its capability areas until all the capability areas are assessed. The capability areas for assessment are presented per domain component in tabular form, with the name of the capability area on the left, followed by a key of its capability category, and its progressive maturity descriptions from *Initial* to *Optimising* next to that. All domain components follow the same format. The data collection assessment sheet can be seen in Figures 5.10 to 5.12. The maturity level descriptions of the capability areas of all the domain components can be seen in Appendix E.

Start assessment | Back to instructions

Organisational level

Key
Primary capability activities
Supporting structures
Enabling practices

Collection	Storage	Sharing	Analysis	Usage	Governance
Validate captured data	Data aggregation and storage for management and decision-making	Data availability between the organisation and the facilities (organisation interoperability)	Data aggregation for analysis	Usage of data for coordinated care	Data strategy
Processing of collected data	Storage of workforce and human capital data	Data availability between the organisation and other providers (provider interoperability)	Preprocessing for analysis	Improve quality and efficiency of patient care	Master data management
Workforce and human capital data collection	Data ingestion	Data processing for data sharing	Storage for analysis	Usage of analysed data	Metadata management
Data collection technology	Data storage backups	Data transmission networks	Data analysis	Data usage for organisation performance reporting	Standards, policies and regulations for primary activities
Data collection infrastructure	Data storage technology	Technology for data sharing with the facilities and other care providers	Data analysis technology, tools and techniques	Oversight of healthcare delivery system	Standards, policies and regulations for supportive structures
Data entry forms/structure	Data storage infrastructure	Infrastructure for data sharing with the facilities and other care providers	Data analysis infrastructure	Population health surveillance	Data life cycle management
Comprehensiveness of data captured	Data quality processes for stored data	Data sharing privacy and security	Data analysis privacy and security	Data usage for workforce and human capital management	Data security and privacy management
Types of data captured	Retention and disposal plan of stored data	Security of data sharing devices	Alignment to data analysis standards, policies and regulations	Data queries for various usages	Adherence to national/governmental healthcare data management acts/regulations/policies
Quality of collected data	Storage privacy and security	Alignment to data sharing standards, policies and regulations	Security of data analysis devices	Data usage technologies	Data business rules
Data collection privacy and security	Alignment to data storing standards, policies and regulations			Data usage infrastructure	Health indicators data collection specification
Security of physical data collection devices				Data usage privacy and security	
Alignment to data collection standards, policies and regulations				Security of data usage devices	
				Data usage alignment to standards, policies and regulations	

Figure 5.8: Organisational-level overview

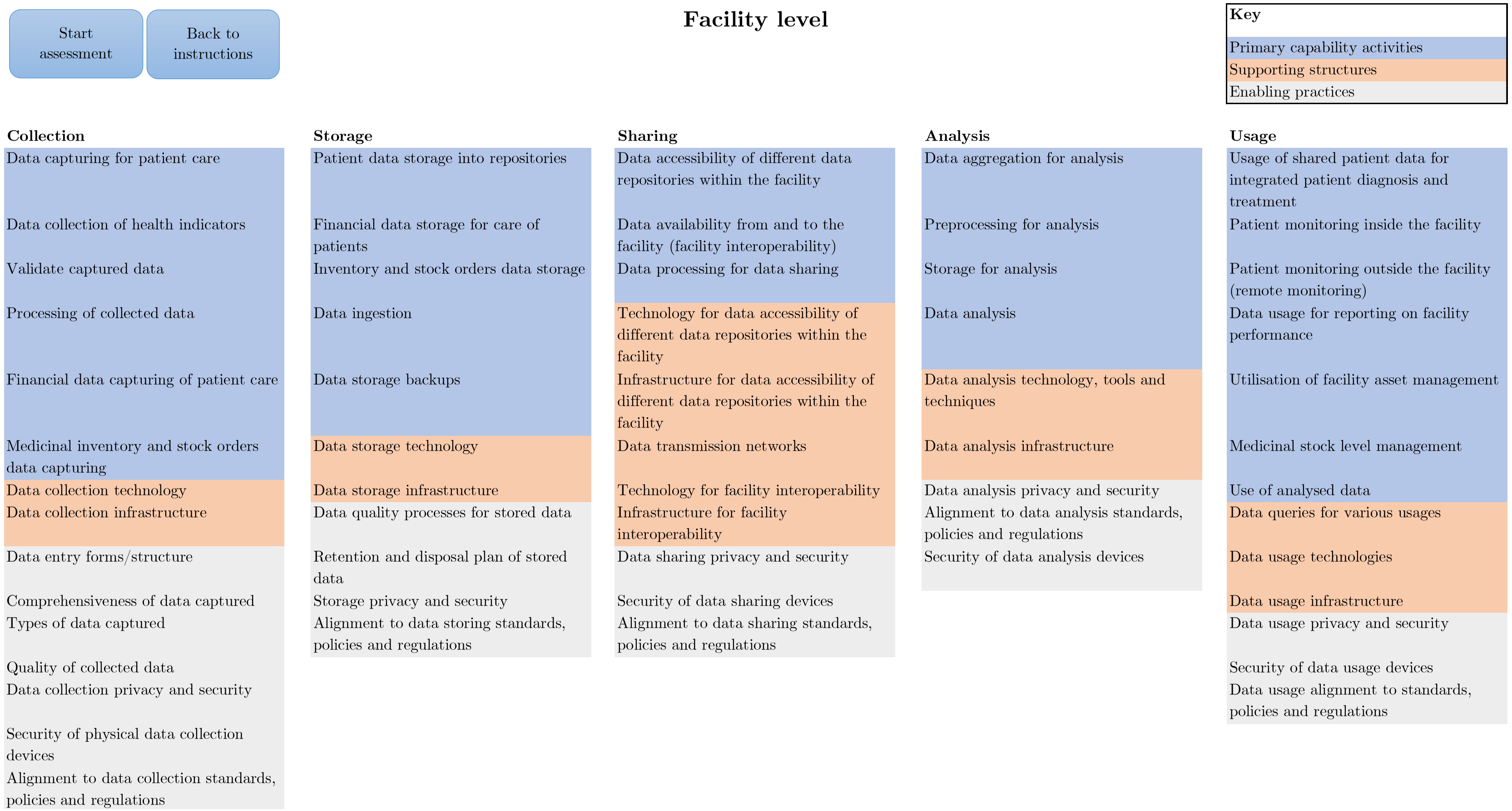


Figure 5.9: Facility-level overview

Data collection

Results

Capability area		1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising	Score
Validate captured data	**PA**	Captured data is not validated	The captured data is validated repeatedly, resulting in the elimination of errors in the captured data. Data validation is controlled through basic methods and procedures. Validation of captured data is time-consuming, is dependent on dedicated individuals and seldom results in validated data that eliminates errors like missing data, data duplication and other inconsistencies effectively	All the different kinds of data validation methods and procedures are well-defined and implemented for effective data validation. The inputs, standards and procedures for validating data are well-defined and implemented. Data validation results in minimal errors in the data	Data validation methods and processes that do not result in effectively validated data are detected and refined to further improve data validation	The range of ability, methods and procedures of the validation of captured data are continuously improved by innovatively exploiting data validation best practices. Data validation continuously improves the automatic data validation of all captured data	
Processing of collected data	**PA**	Captured data is not processed to ensure it is useable	Basic captured data processing methods and procedures are repeatable. Raw data is transformed to be useable and easily interpreted, but the processing is resource-intensive and time-consuming	The data processing methods and procedures are well documented. The processing methods and procedures are effective at making data easily understandable and useable. Processing is standardised and consistent	Data processing methods and processes that do not result in easily understandable and useable data are detected and refined through corrective measures that are in place	The processing of data is continuously improved by innovatively exploiting best data processing practices. Collected data processing continuously improves the automatic data processing of all captured data	
Workforce and human capital data collection	**PA**	No clear procedures and processes exist to collect workforce and human capital data. Data is seldom collected. Collected data is disorganised	Workforce and human capital data is collected following basic defined processes and procedures. Data is repeatedly collected and collection is organised. Different facilities' data are in silos. A basic set of workforce and human capital data is collected	Workforce and human capital data is collected following standardised procedures and processes across the whole organisation. Collected data is consolidated. The data that needs to be collected is standardised across all organisational entities. Workforce and human capital data is collected for specific purposes	The execution of the collection of workforce and human capital data is supervised and monitored. Areas of ineffective data collection can be detected and corrective action is taken. The skills, knowledge and experience of the workforce are collected	What workforce and human capital data to collect and the collection procedures and processes are continuously optimised in order to determine indicators like the supply, demand and distribution of the workforce	

Figure 5.10: Assessment sheet of data collection on the organisational-level

Data collection technology	**SS**	Technologies for data collection exist. The available technologies give very limited support to accomplish effective data collection	Data collection technologies are adequate, robust, reliable and available. Adequate hardware and software are available for repeatable data collection	Data collection technology is compatible and interoperates with other devices and applications. Technologies for structured and unstructured electronic data collection are applied	Availability, reliability and maintainability of data collection technology are monitored and corrective maintenance is executed effectively and timely when necessary	Preventive maintenance and upgrades of data collection technology are executed proactively, timely and effectively. New innovative technologies are continuously implemented	
Data collection infrastructure	**SS**	Infrastructure for data collection exists, but gives very limited support to the effective execution of data collection	Appropriate infrastructure for data collection is applied that gives adequate support for repeatable data collection	Specific infrastructure for data collection is applied that supports data collection effectively. Infrastructure for structured and unstructured data collection are applied	The availability, reliability and maintainability of data collection infrastructure are monitored. When failure occurs infrastructure is repaired or replaced	Data collection infrastructure is continuously maintained and upgraded whenever needed. Infrastructure is continuously improved towards new innovative or alternative structures	
Data entry forms/structure	**EP**	Initial data entry forms do not enable the capturing of all relevant data. Entry forms are not easy to use and important data is left out	Data entry forms enable the adequate and consistent capturing of relevant data.	Well-established data entry forms enable the effective capturing of all relevant data. The data entry forms are easy to use, intuitive and not time-consuming. The data entry forms are standardised across the different points of care	The data entry forms are controlled and checks are in place to ensure appropriate data entry onto the forms across the different care points	The data entry forms for data capturing are continuously optimised to enable the most effective data capturing. The ease to enter data and time it takes are continuously improved. Measures to ensure data is correctly entered are continuously improved	
Comprehensiveness of data captured	**EP**	Data is being collected without the adequate knowledge of what all the necessary data elements are for baseline care delivery and decision-making	Data is being collected with the adequate knowledge of what all the necessary data elements are for baseline care delivery and decision-making	Comprehensive data is being collected, according to the extensive knowledge of what all the necessary data elements are, for effective care delivery and decision-making	It is monitored whether the comprehensive sets of data are effectively applied according to the purpose that they were collected for. Corrective action is taken when data is not effectively applied or collection of data is discontinued if redundant	The comprehensiveness of data is continuously improved when new applications for data are not being collected yet surface for optimised care delivery and decision-making	

Figure 5.11: Assessment sheet of data collection continued

Types of data captured	**EP**	Paper-based and electronic data are captured. Data is not captured in a controlled way. Vast amounts of data are captured on paper	Paper-based and electronic data is captured. Electronic data collection consists of structured data only	Paper-based and electronic data structures are defined. Electronic data types that are collected include structured and unstructured data	Paper-based, structured and unstructured electronic data capturing are managed effectively. If paper-based capturing is done, it is because it is more effective than electronic capturing for its purpose and not because of legacy systems	The various types of paper-based and electronic data that are captured are continuously improved for the efficient and effective application of the collected data according to its purpose	
Quality of collected data	**EP**	Collected data is of varying and unacceptable quality	Data of similar quality is consistently collected, but no defined standard of quality data exist	The designated role and purpose, consistency and timeliness of collected data are defined and standardised. An entity-wide understanding of the quality of data metrics exists	Data is collected effectively, consistently and timely according to its designated role and purpose	Causes of unacceptable quality data are continuously and effectively identified and addressed	
Data collection privacy and security	**EP**	Collected data is not secured and patient privacy is not guaranteed	Basic authorised access control to collected data exists to ensure the availability, integrity and confidentiality of healthcare data. Access options between view data, insert data, update data and delete data is differentiated for privacy and security	Security software is in place to protect collected data against malware and procedures are in place to prevent malicious use of data by authorised users. Operational and management security for data collection are well-defined.	Confidentiality and integrity of collected data are monitored to ensure data stays available to only the intended users and to ensure data is not corrupted or stolen. Measures are in place to identify malicious use of data through authorised users	Continuously upgrade and update data collection privacy and security software. Security and privacy procedures and checks are continuously improved to ensure availability, integrity and confidentiality of the collected data	
Security of physical data collection devices	**EP**	Data collection devices are not secured. Collection devices can easily be stolen, preventing effective collection of data	Basic security policies and procedures are in place for the security of physical data collection devices. Security policies and procedures do not completely ensure devices are secure and not stolen	Security policies and procedures for physical data collection devices are well defined and established, ensuring a stable environment for data collection through the devices	The security of physical data collection devices is monitored. The correct usage of the devices is monitored to ensure they are not used for the wrong intent. Users that use the devices for the wrong reasons can be identified	The security of physical data collection devices is continuously improved. Device security policies and methods are continuously updated to prevent new threats and misuses of the devices	

Figure 5.12: Assessment sheet of data collection continued

Once the assessment is completed, the user can navigate to the results sheet by clicking on the results button. The results sheet with hypothetical results are presented in Appendix F. The results sheet include an explanation of the presented results, along with indicators of how the results should be interpreted and a table with an overview of the results per domain component and capability category. The results sheet also includes visual representations of the overall results. Lastly, a table is included with the scores of all the capability areas which can be filtered according to different criteria in order to customise the user's view of the results. A visual representation of the customised views is also available.

5.6 Reflection on future upgrades

Future upgrades of the HCDMMM are of key importance as the application environment of the HCDMMM changes. As time goes by, components of the HCDMMM will become obsolete, new constructs will emerge that will be needed to be included in the HCDMMM and the requirements to reach different maturity levels might also change. Therefore, it is important to reflect on how upgrades will be introduced to the HCDMMM. The reflection on future upgrades can be seen in Table 5.8. This is step five of the design decisions that Mettler (2010*b*) proposed.

Table 5.8: Reflect future upgrades phase decisions (Mettler, 2010*b*)

Design parameter	Chosen characteristic	Decision description
Subject of change	None	The HCDMMM does not comprise any formal mechanisms for changing the HCDMMM, but the different components of the HCDMMM will be revised based on users' inputs
Frequency	Non-recurrent	Changes to the HCDMMM will not occur at predetermined intervals, but when necessary
Structure of change	Internal / exclusive	Changes to the HCDMMM can be done exclusively by the developer, but users can make their own changes for their specific usage
Dissemination	Open	The HCDMMM is freely available upon request from the developer

It was decided that there will not be a specific subject of change (as opposed to a subject of change of form, functioning or form and functioning), but as time goes by and changes to the different components of the HCDMMM are necessary, they will be introduced. Users' input will also be a driver of change. Not specifying a subject of change means that all components of the HCDMMM can be changed when it is necessary.

The frequency of changes will not happen continuously or at predetermined intervals. Rather, they will be introduced to the HCDMMM when necessary due to components that became obsolete, the emergence of new constructs or the change in requirements to reach maturity levels. This decision enables changes to the HCDMMM whenever it is necessary.

Furthermore, the changes to the HCDMMM can be implemented by the HCDMMM developer (as opposed to externally by other users). This restricts who can make changes to the HCDMMM and enables control over which changes are introduced to the HCDMMM. Lastly, the HCDMMM is freely available upon request from the developer so that anyone who believes they can benefit from using the HCDMMM can have access to it (as opposed to it being available to an exclusive group of users).

5.7 Conclusion on the HCDMMM development and presentation chapter

In this chapter the development process of the HCDMMM was described and the final HCDMMM was presented. This chapter starts with an outline of the development methodology which was derived from literature as presented in Chapter 4. This literature consisted of defined development methodologies as described by De Bruin *et al.* (2005) and Becker *et al.* (2009), design decisions by Mettler (2010*b*) and design principles by Pöppelbuß and Röglinger (2011).

Subsequently, the chapter described how the development methodology was executed. The first step of the development methodology was to describe the need for a new maturity model as determined from the preceding chapters. Chapters 2 and 3 identified the need for a research product that assists in addressing the data management challenges in healthcare and in Chapter 4, the need for another maturity model was also justified.

The next step was to define the scope of the maturity model. It set the boundary of what the HCDMMM focuses on, which components fell in and out of the scope of the HCDMMM and who the HCDMMM is developed for.

Next, the actual development of the HCDMMM was described. Design phase decisions according to Mettler (2010*b*) were first explained, followed by the iterative development process of constructing the HCDMMM architecture, populating the HCDMMM and verifying it through various methods which included SMEs' contributions. The development of the different components was described individually and the final result of each model component was presented.

Subsequently, the complete HCDMMM was presented with all its different components. It was first presented conceptually and then as it will be transferred to users and will be used by them in practise.

The last section of this chapter was the reflection on future upgrades of the HCDMMM. Due to the fact that the environment which the HCDMMM will be used in will change over time, it was necessary to determine how changes will be introduced to the HCDMMM in the future to keep it updated and prevent it from becoming obsolete.

The HCDMMM was subjected to evaluation during and after the development process. Therefore, the next chapter is dedicated to describing how the evaluation of the HCDMMM was conducted. The evaluation consisted of verification and validation. These two processes are explained in Chapter 6.

Chapter 6

Evaluation of the HCDMMM

In this chapter, the evaluation of the developed Healthcare Data Management Maturity Model (HCDMMM) is presented. The two processes followed to evaluate the developed model include (i) verification and (ii) validation. Verification was done to ensure that the HCDMMM is feasible, theoretically sound and adheres to the specified requirements. The validation of the HCDMMM illustrated that it addresses the defined problem of this study and that it is useful to the intended users. Furthermore, the design decisions outlined by Mettler (2010*b*) were considered during the development of the evaluation strategy. The rest of this chapter outlines the evaluation strategy (Section 6.1), the results from the verification stage (Section 6.2) and the results from the validation stage (Section 6.3). Lastly, this chapter is concluded in Section 6.4.

6.1 Evaluation strategy

The evaluation strategy consisted of verification and validation stages. The purpose of the verification process was to determine whether the HCDMMM addressed all the formulated design requirements, covered all the key elements of relevance and defined the key elements and their maturity correctly. The purpose of the validation process was to determine whether the HCDMMM could be deemed to be usable and achieved its intended purpose. The evaluation strategy is depicted in Figure 6.1.

Verification determines whether the developed model represents the developer's conceptual description and specifications with sufficient accuracy (Conwell *et al.*, 2000). This ensures that the model is theoretically sound and that it is feasible.

Validation ensures that the model that was developed according to certain specifications is aligned with the needs of the user. Validation determines to what degree the maturity model is an accurate representation of the real world from the perspective of its intended user (Conwell *et al.*, 2000).

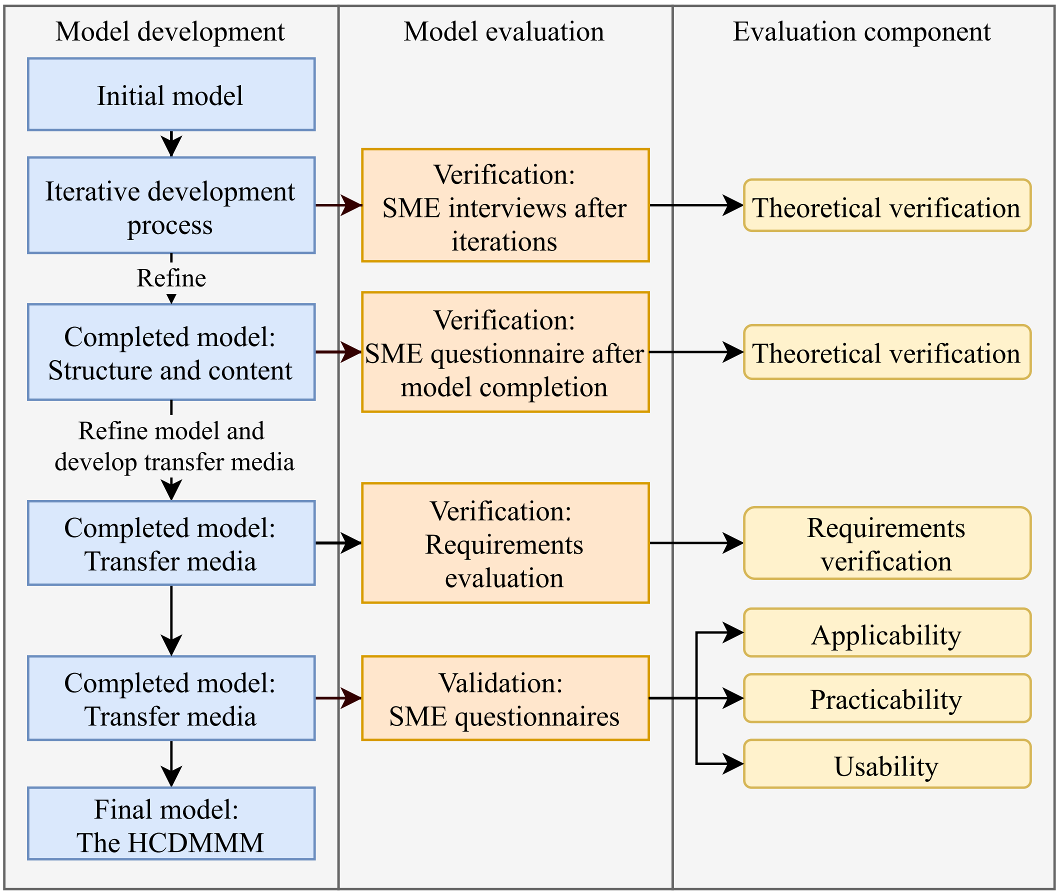


Figure 6.1: Evaluation strategy

The design decisions of the evaluation phase as discussed by Mettler (2010*b*) were incorporated in the evaluation strategy. The relevant design decisions for this phase are illustrated in Table 6.1. The overarching evaluation strategy is described first, followed by the verification strategy and lastly, the validation strategy.

The first design decision for the evaluation phase was to determine whether the *subject of evaluation* is the design product, the design process or both of these. It was decided to focus on the evaluation of the design product. By evaluating the design product, it ensures that the developed model is theoretically sound and usable by the intended users. Although a rigorously defined methodology was used, incorporating design decisions and design principles to ensure that a sound design process was followed that yielded an accurate and usable model, this design process was not evaluated.

For the *point of time* evaluation decision, it was decided that the evaluation of the HCDMMM would happen both ex-ante and ex-post. The verification of the HCDMMM happened during the development of the HCDMMM, because this ensured that components that are further developed, are built on already verified components. Based on the verification feedback, the HCDMMM was refined to address the issues highlighted in the feedback. Verification also happened after the development process to verify that the completed HCDMMM was theoretically sound and addressed the specified requirements. The validation of the HCDMMM happened after it was completed, in order to evaluate the HCDMMM after the refinements from the verification process were incorporated.

The last evaluation decision made was that a naturalistic evaluation method would be followed, rather than an artificial evaluation method. Given that a naturalistic evaluation method was decided on, the HCDMMM was evaluated through interviews with intended users. SMEs verified that the HCDMMM as developed was theoretically sound, and validated that it is applicable, practicable and usable. Intended users also gave their opinion on the strengths and weaknesses of the HCDMMM.

Table 6.1: Evaluate phase decisions (Mettler, 2010*b*)

Design parameter	**Chosen characteristic**	**Verification**	**Validation**
Subject of evaluation	Design product	Determine whether the design product built was theoretically sound and that it met the requirements	Determine whether the design product is applicable to real-world cases, practicable and usable
Point of time	Ex-ante and ex-post	Happened during and after the development process	Happened after the HCDMMM was developed
Evaluation method	Naturalistic	Knowledge domain SMEs were interviewed to determine whether the HCDMMM is theoretically sound, feasible and meets the specified requirements	Intended users are interviewed to determine whether the HCDMMM achieves its intended purpose, is usable and practicable to the intended user and is applicable to real world cases

The remainder of this section expands on the evaluation strategy. It consists of the explanation of the verification strategy (Section 6.1.1) and the validation strategy (Section 6.1.2).

6.1.1 Verification strategy

In accordance with the description of verification that was given at the beginning of this section, the verification strategy that was followed for this study is described. The verification strategy consists of the theoretical verification process and the requirements specifications evaluation.

The theoretical verification consisted out of two verification stages. The one stage took place during the development process. As components were developed, they were also verified throughout the development process in order to determine whether refinement is necessary. The other verification stage took place when the development of the HCDMMM content was completed. When the HCDMMM content was completed, the overall HCDMMM was verified to evaluate all of its aspects comprehensively. This verification stage yielded some final refinements before the evaluation of the requirements specification and the validation of the HCDMMM.

The requirements for the proposed research product for this study are outlined in Chapter 3. In order to ensure that the developed HCDMMM addressed the specified problem, it had to address the specified requirements. Therefore, the verification process evaluated that the developed HCDMMM addressed the specified research product requirements of Chapter 3 with sufficient accuracy.

The sections that follow describe the verification strategy in more detail. In Section 6.1.1.1 the theoretical verification strategy is described and in Section 6.1.1.2 the requirements verification strategy is described.

6.1.1.1 Theoretical verification strategy

The theoretical verification strategy consisted of two stages, the first stage being the verification of components during the development process and the second stage being the verification of the complete HCDMMM evaluating all of its aspects comprehensively. For the first verification stage, it was decided that as the components were developed, they were to be verified through the assistance of SMEs in various knowledge domains. This verification process was to consist of semi-structured interviews with the SMEs from different knowledge domains. Semi-structured interviews were to be held with different SMEs at specific iterations of the development process. The reason for this was to verify components and content as they were developed. SMEs from different knowledge domains were to be included to ensure the HCDMMM was verified from all perspectives that were incorporated in the HCDMMM. The knowledge domains were to include maturity models, SQL, enterprise architecture, and data engineering. The SMEs were to be introduced to the development process at specific iterations and in a specific sequence that would ensure that the HCDMMM was being built on already verified components. The purpose of this verification process during the development of the HCDMMM was to establish that the components that were developed were theoretically sound and feasible, to make changes to components that did not sufficiently represent the requirements or to add components that ensured the accurate representation of the specified requirements.

After the content of the HCDMMM was completely developed with the assistance of SMEs, it was also necessary to determine whether the overall developed HCDMMM represented the specified requirements with sufficient accuracy. It was decided that during this phase of the verification process SMEs in healthcare data management would be used to test whether the HCDMMM is theoretically sound. Using SMEs in healthcare data management ensured that they are capable of evaluating the HCDMMM comprehensively. It was also decided that these SMEs would complete a questionnaire to verify the HCDMMM comprehensively. The attention points were not verified through the SMEs as these points are only desirable, should be noted and considered, but do not have to be met as a mandate.

6.1.1.2 Requirements specification evaluation strategy

Finally, tables were to be constructed according to each requirements category that identified which components of the HCDMMM satisfied which requirements. This served as an indication that the HCDMMM addressed all the requirements, as well as which components of the HCDMMM fulfilled which requirements. The fulfilment of these requirements is presented in Appendix H.

6.1.2 Validation strategy

The validation strategy for this study is described in accordance to the description of validation that was given at the beginning of this section. Validation determines whether the developed model addresses the specified problem, whether it is an accurate representation of the real world from the perspective of the intended users of the model and whether it is aligned with the needs of the user. Therefore, to validate the HCDMMM, it was assessed along the dimensions of:

- applicability to real-world cases: the conceptual structure and components of the HCDMMM are representative of the data management of real-world healthcare entities on a facility and an organisational level;
- practicability: the HCDMMM and its assessment methodology can be put into practice; and
- usability: the degree to which the HCDMMM can be used by the intended user.

The HCDMMM validation was conducted by means of a questionnaire that consists of different sets of questions and statements, with each set of questions focused on validating a different component of the validation strategy. The respondents could specify on a scale how strongly they agree with every question or statement. By considering all the allocated scores to the different components of the HCDMMM, the HCDMMM was validated comprehensively. The SMEs were also required to give their opinion on the strengths, weaknesses and where the HCDMMM were to fail with regard to each validation dimension.

The respondents of the validation strategy were determined to be SMEs in management positions of healthcare entities with knowledge of data management processes on a facility level and an organisational level. The focus was on developing a model that can help improve data management in the public sector in developing countries, but during the validation process it is also determined whether the HCDMMM can be used in the private sector of a developing country too. As this study was undertaken in South Africa, it was decided to mostly use SMEs of healthcare entities in South Africa.

6.2 Verification

The verification of the HCDMMM was executed according to the verification strategy as described in Section 6.1.1. It consisted of theoretical verification (Section 6.2.1), that was done during and after the development process, and the evaluation of the requirements specifications (Section 6.2.2).

6.2.1 Theoretical verification

This section describes the execution of the theoretical verification. It consists of two stages, which are verification throughout the development process (Section 6.2.1.1) and verification after the development process (Section 6.2.1.2). The last part of this section describes the HCDMMM refinements due to verification after the development process (Section 6.2.1.3).

6.2.1.1 Verification throughout the development process

This section describes how the verification strategy throughout the development process was executed. The HCDMMM design and content were verified throughout the development process by means of semi-structured interviews with SMEs. Questions were constructed to verify the different structures and content of the HCDMMM both on a conceptual and more detailed level. Questions were customised for each SME according to the SMEs' knowledge domain and according to the components they verified. SMEs that were used had knowledge of SQL, maturity models, enterprise architecture, and data engineering. An iterative development process was followed and different SMEs were introduced at different stages of the development process to verify the HCDMMM. As components were near completion, SMEs were consulted to ensure those components were built theoretically sound, and if not, necessary changes were made. The SMEs were introduced to the development process at specific times and in a specific order to verify components sequentially. This ensured that further development of the HCDMMM was built on already verified components.

The SME on maturity models verified the maturity progression of the capability areas. The SME on SQL and enterprise architecture verified the transactional information system's components of the healthcare data management system, and the different capability area categories, namely primary activities, supporting structures and enabling practices. The data engineering SMEs verified the different data management concepts that many capability areas were based on. They also verified many individual capability areas and their descriptions. The data engineering SMEs had many valuable insights on all the capability areas across the domain components as they had knowledge of all the necessary primary activities, the technological and infrastructural components and the regulatory environment of data management. Table 6.2 illustrates the knowledge domains and professional status of the different SMEs.

Table 6.2: Verification SMEs

SME	**Relevant knowledge areas**	**Professional status**
SME 1	Maturity models	PhD in Industrial Engineering where he developed a capability maturity model focused on pharmacovigilance
SME 2	Enterprise architecture, SQL	Project manager and technical specification expert on information system implementation projects
SME 3	Data Engineering	Senior data engineer
SME 4	Data Engineering	Senior director of data systems

Appendix G contains the summaries of all the interviews held with the verification SMEs. The parts they verified are stated there, as well as the refinements suggested to the HCDMMM. Appendix D outlines all the refinements that occurred throughout the different iterations of the HCDMMM development and which of the SMEs had influence on the changes at the different iterations.

6.2.1.2 Verification after the development process

The verification step following the development process was carried out after the structure and content of the HCDMMM was completed. This means the conceptual HCDMMM as illustrated in Section 5.5 was verified before the transfer media of the HCDMMM (illustrated in Section 5.5.2) was developed.

The primary focus of this verification step was to indicate whether the model development through the use of SMEs' inputs resulted in a model with a structure and content that addresses the specifications. This verification step also yielded some refinements to the HCDMMM. This verification step consisted of correspondences with healthcare data management SMEs. The correspondence included a questionnaire with the purpose of determining how well the different requirements were addressed by the developed HCDMMM, as well as a semi-structured interview that focused on how to refine the HCDMMM further.

Two healthcare data management SMEs were consulted and requested to complete a questionnaire (Shown in Appendix G.4) that verified that all the components of the HCDMMM met the requirements. The reason why healthcare data management SMEs were included was because of their comprehensive knowledge of data management, specifically in the context of healthcare, which enabled them to verify the HCDMMM comprehensively. The profiles of the healthcare data management SMEs is given in Table 6.3.

Table 6.3: Healthcare data management verification SMEs

SME	**Relevant knowledge areas**	**Professional status**
SME 1	Data, healthcare	Data scientist at private healthcare group in South Africa
SME 2	Data, healthcare	Process improvement engineer for private healthcare group in South Africa

The questionnaire employed a five-point scale to determine how well each requirement was met. Requirements scoring an average of 4 or 5 were deemed as having addressed the specifications satisfactorily and no further refinements were needed, but small refinements could improve the HCDMMM further. Requirements with an average score of 3 were considered as uncertain whether the specifications were addressed satisfactorily and further refinements were likely needed. Lastly, requirements that received a score of 1 or 2 were considered as not having addressed the specifications with sufficient accuracy and further refinements were needed.

The scores of the two healthcare data management SME are given in Table 6.4 per target requirement. The questionnaire completed by the two SMEs and their accompanied comments can be viewed in Appendix G.4.

Table 6.4 provides the average score per target requirement. The scores of the questionnaire questions that verified the different components are discussed during this section. Follow-up interviews were also held with the SMEs to discuss further model refinements.

Table 6.4: Healthcare data SME responses

Target	**Verification question or statement**	**SME 1**	**SME 2**	**Avg**
FR1	The use of the HCDMMM to assess an entity's healthcare data management leads to the eventual improvement of better care delivery to patients and better care delivery management and decision-making	4	4	4
FR2	The HCDMMM can be used to make an as-is assessment of an entity's healthcare data management	5	3	4
FR3	Do the maturity levels accumulate (improve incrementally), with each level encompassing the preceding lower level of maturity?	4	2	3
FR4	To what extent do you agree with the capability areas included under data collection?	4	4	4
FR4	To what extent do you agree with the capability areas included under data storage?	4	4	4
FR4	To what extent do you agree with the capability areas included under data sharing?	4	4	4
FR4	To what extent do you agree with the capability areas included under data analysis?	4	4	4
FR4	To what extent do you agree with the capability areas included under data usage?	4	4	4
FR4	To what extent do you agree with the capability areas included under data governance?	4	4	4
FR5	To what extent do you agree with the two system levels, the facility- and organisational-level, that are included?	5	4	4,5
FR6	To what extent do you agree with the supporting structures included in the HCDMMM?	4	4	4
UR1	The HCDMMM can be used by managers or change agents in healthcare entities to assess the maturity of the entity's data management	4	2	3
UR2	To what extent do you agree that the HCDMMM is generic so that it can be used by more than one entity?	4	4	4

Continued on next page

Continued from previous page

Target	**Verification question or statement**	**SME 1**	**SME 2**	**Avg**
UR3	The HCDMMM is user friendly and intuitive	3	1	2
UR4	The HCDMMM uses standard domain language and is therefore, easily understandable	3	3	3
BC1	The HCDMMM considers all the necessary governmental and national policies, acts and regulations	3	1	2
BC2	The HCDMMM considers ethical considerations with regard to data management in healthcare	4	4	4
BC3	The HCDMMM includes the privacy and security of data	4	4	4
BC4	The HCDMMM includes all the necessary internal standards and policies that an entity should consider	4	2	3
DR1	The HCDMMM is not applicable to other domains other than healthcare data management	1	2	1,5
DR2	The HCDMMM does not specify any specific technology and infrastructure that should be incorporated to achieve a maturity level	5	5	5
DR3	The HCDMMM describes different maturity levels of capability areas without prescribing how to achieve maturity levels	5	5	5

The results of the questionnaire were mostly positive with some degree of uncertainty and a few components that SMEs responded to with 'strongly disagree' and 'disagree'. Target requirements that were verified as having sufficient accuracy are FR1, FR4, FR5, FR6, UR2, BC2, BC3, DR2 and DR3. Taking both SMEs' scores into consideration, there was still some uncertainty about requirements FR2, FR3, UR1, UR3, UR4 and BC4. Lastly, the two requirements that seemingly were not addressed are BC1 and DR1.

The requirements that there was uncertainty over and that were not met thus far by the HCDMMM are discussed further. As this verification step was conducted before the transfer media was developed, as described in Section 5.5.2, it explains the uncertainty of FR2, UR1, UR3 and UR4. After the transfer media was completed and the final refinements were made to the completed HCDMMM, these requirements were also addressed satisfactorily.

FR2: Although SME 1 gave this requirement a score of 5 commenting "very useful in benchmarking of a healthcare entity's healthcare data management.

I have not come across a tool or model that creates this capability as well as this model does", SME 2 was unsure whether the HCDMMM is practical to make an as-is assessment as it "left room for interpretation by assessor" due to words used that are unscientific and unquantifiable which is also the reason for the weak score given to FR3. SME 2 suggested that this part of the HCDMMM should be evaluated further through validation. Taking both opinions into consideration, the HCDMMM would be practical to make an as-is assessment when there is better distinction between levels through more quantifiable and concise terminology. After level descriptions were improved, as well as developing the transfer media of the HCDMMM to make assessments easier and practical, this requirement was further evaluated through the validation phase.

FR3: SME 1 agreed that the levels accumulate and improve incrementally, and commented that in practice some of the incremental levels might be more exponential than linear and "would require significantly more effort to move between maturity levels." SME 2 felt that due to vague terminology, the lines between different maturity levels become blurry. Taking both SMEs' opinions into consideration, it seems that there is good incremental progression between levels, but terminology used like "reliable", "appropriate" and "available" that are not very specific make it difficult to quantify on which level an entity is and makes an objective, quantified assessment of capability areas difficult. By addressing this, the uncertainty of FR2 was simultaneously addressed too.

UR1: SME 1 agreed that this requirement is fulfilled, but stated that definitions of the HCDMMM need to be agreed upon, as some definitions are open for interpretation. Due to the same reason (blurry maturity levels) for the weak score of FR3, SME 2 also scored this requirement low commenting "I question the practicality of the model." After the improvement of maturity level descriptions and the development of the transfer media that improved the practicality of the HCDMMM significantly, this requirement was further evaluated during validation.

UR3: This requirement also received a low score. SME 1 commented that training would be required by HCDMMM users and there was uncertainty over how users were supposed to use the HCDMMM to make an assessment. This score was given before the transfer media was developed and after it was completed and shown to SME 1 during a semi-structured interview, he agreed that the transfer media gives sufficient information and instructions to the user to know how to use the HCDMMM and the uncertainty as to how assessment will be done was ruled out. SME 1 also stated that the HCDMMM transfer media provided a progression and flow through the assessment which makes it a lot more user-friendly and intuitive to use. Nevertheless, this requirement was further evaluated during validation. SME 2 strongly disagreed for the same reason as FR2 and FR3 and is of the opinion that the HCDMMM will be user-friendly and intuitive if it provides an incontestable method of maturity assessment. This was addressed by improving the level descriptions further and the development of the transfer media with assessment instructions, which improved the user friendliness and intuitiveness of the HCDMMM.

UR4: Both SMEs were unsure whether the HCDMMM used standard domain language. SME 1 said the score given was due to standard domain language being different for management, healthcare professionals or technical staff. After elaborating during a follow-up interview that it was intended for management, he agreed that the HCDMMM incorporated standard language, but advised that it should establish definitions for different concepts so that all the users have the same reference for different terminology. This was included in the transfer media. SME 2 said standard domain language was used, but was not concise enough.

BC4: SME 1 agreed that this requirement was fulfilled and that "the model has nothing missing that is apparent". Although SME 2 agreed that standards and policies are mentioned, he disagreed that this requirement was fulfilled, commenting that no reference was made to specific standards and policies. The HCDMMM was developed to be generic so that it can be used by more than one entity (UR2), it was intentional that no specific standards and policies are referred to. SME 1 also commented that internal standards and policies are highly specific to a given entity and referring to specific standards and policies will cause the HCDMMM to lose its generic applicability. Therefore, the HCDMMM includes a generic assessment of the internal standards so that it can be applicable to different entities.

Both SME 1 and SME 2 seemingly disagreed that the following two requirements that are discussed next were not fulfilled. These two requirements are DR1 and BC1.

DR1: Both SME 1 and SME 2 felt that the HCDMMM can be applied to other data management domains, even though this requirement states that it should only be applicable to the healthcare sector. SME 1 gave it a strongly disagree score, even though he stated that the HCDMMM is specifically very applicable to healthcare. The reason he gave it a strongly disagree score, was that the baseline of the HCDMMM could be applied to other data management domains with some required adjustments. After a further interview with SME 1, this was taken as positive feedback as it enlarges the scope of applicability of the HCDMMM, while it fulfils the requirement of developing a data management maturity model specifically for the healthcare sector.

BC1: SME 1 scored this requirement as unsure and commented that he was not qualified to comment on this. SME 2 strongly disagreed that this requirement is fulfilled and commented that no references were made to specific policies, acts and regulations. As stated earlier, another requirement of the HCDMMM is to be generic enough so that it is applicable to different entities (UR2) and even to different countries (AP1) and that is why no reference was made to specific governmental and national policies, acts and standards, as it will differ for different contexts, but these policies, acts and regulations are referred to in general so that entities can make an assessment of that capability area. Two other SMEs also verified this during the design process. They both referred to the Protection of Personal Information Act (POPIA) with regard to the South African context as an Act that is very relevant to healthcare data management and needs to be addressed by the HCDMMM. The capability area "Adherence to national/governmental healthcare data management acts/regulations/policies" addresses acts like these without referring to specific ones to maintain its generic nature, thus fulfilling the requirement.

6.2.1.3 Refinement of the HCDMMM due to verification

During the verification of the HCDMMM that took place after the development process, it was verified that the HCDMMM was developed with sufficient accuracy, but some components were still lacking and not all requirements were adequately addressed. Therefore, additional refinements were made to the HCDMMM to ensure all the requirements were addressed. These refinements included: (i) the development of the transfer media, as discussed in Chapter 5.5; (ii) the refinement of all capability level descriptions; (iii) the clarification of vague terminology and the addition of concept definitions; and (iv) the addition of some capability areas.

The transfer media, as illustrated in Chapter 5.5, was developed to improve the HCDMMM's practicality for making an as-is assessment and to improve the ease-of-use of the intended users. Also incorporated into the transfer media of the HCDMMM to improve the HCDMMM's ability to make as-is assessments and its usability, was introductory information and instructions on how to use the tool to make assessments, concept definitions and terminology clarifications that eliminate ambiguity. The introductory information on the HCDMMM gives information on the use of the HCDMMM so that training is not needed. The instructions are also clear so that all users can assess maturity in the same way, thus making it intuitive to use. The instructions on how to make an assessment also provides a way to make consistent assessments and conveys the incontestable method of making an assessment. The navigation pages of the transfer media makes it easy to move between different pages and the results sheet makes it easy to scrutinise the results of different capability areas and thus, making it user-friendly.

The level descriptions of all capability areas were refined to make them more concise and specific. This ensured that different maturity levels are not blurry and were thus distinct and progressive. Much vague terminology was substituted with quantifiable terms, or terms that were not substituted were clarified so that all users have the same understanding of the terms. This refinement also improved the HCDMMM's practicability and applicability to make as-is assessments and its usability, so that all users will execute assessments in the same way, as level descriptions are clearer, maturity levels are distinct and progressive and all users have the same understanding of different terminology and concepts.

Some capability areas were also added that expanded the HCDMMM further. The addition of these capability areas ensured that data management in the healthcare sector can be assessed more comprehensively by the HCDMMM even though the domain components already addressed the requirements satisfactorily without the additions. After these last refinements, the HCDMMM addressed all the specified requirements. How the HCDMMM addressed all the requirements is described in the next section.

6.2.2 Requirements specification evaluation

This section describes how the requirements that were set and verified, are addressed by the developed HCDMMM. Each requirement of every category was evaluated individually whether it is satisfied by the HCDMMM conceptually or through different components of the HCDMMM. The descriptions of how the HCDMMM addressed every specified requirement is presented in Appendix H. In Appendix H each requirement category is presented in its own table, with all its specified requirements and how the requirements were addressed, to indicate that the HCDMMM satisfied all the specified requirements.

6.3 Validation

This section outlines how validation was executed and presents the validation results. Firstly, how validation was executed is described (Section 6.3.1). A discussion of the validation results follows (Section 6.3.2) and lastly, the refinements that were introduced to the HCDMMM due to validation are described (Section 6.3.3).

6.3.1 Validation execution

Validation was executed according to the validation strategy that is described in Section 6.1.2. The dimensions that the HCDMMM was evaluated against during validation is also specified in Section 6.1.2. Validation can be executed in many ways such as: (i) case studies; (ii) peer reviews; (iii) triangulation; (iv) negative evidence; and (v) member checking. For the purpose of this study a number of SMEs were involved to determine whether the HCDMMM addressed the validation dimensions satisfactorily. A 5-point Likert questionnaire with specific questions and statements was constructed which the SMEs answered in order to validate each dimension. Other studies that also incorporated SMEs through 5-point Likert questionnaires during validation include studies conducted by Schurer (2020), Kleynhans (2020) and Mapowo (2019). Using the approach of incorporating healthcare data management SMEs through a 5-point Likert scale was deemed to be the suitable approach for this study, as healthcare data management SMEs have adequate knowledge to validate whether the HCDMMM will be applicable, practicable and usable in real-world settings. They understand the context that the HCDMMM will be used in and are therefore able to validate the HCDMMM. Using a 5-point Likert scale further improves the appropriateness of executing this approach as it results in a quantitative score of the degree to which the SMEs agree that the HCDMMM is applicable, practicable and usable.

The questionnaire deployed a 5-point Likert scale to measure the three dimensions of validation for this study. This questionnaire can be seen in Appendix I.1. The SMEs were also asked to give their opinion on the strengths and weaknesses of the HCDMMM, and where the HCDMMM might fail with regard to the three validation dimensions. The choices of the Likert scale ranged from strongly agree to strongly disagree to get a holistic view of SMEs' opinions. It also included a mid-point for those who are neutral on the subject matter. All the different statements included in the questionnaire related to a dimension of validation for this study. The combined score of the different statements gave an accurate measurement for the validation of the HCDMMM. The questions included in the validation questionnaire, and the dimensions they targeted, can be viewed in Table 6.5.

A number of SMEs were asked to assist with the validation of the HCDMMM through the completion of the questionnaire. This included SMEs that are employed at healthcare organisations, or organisations that are involved in healthcare strengthening projects, that have comprehensive knowledge on the management of data in the healthcare sector. This included the management of data on a facility level such as hospitals and clinics, as well as the management of data on an organisational level such as the headquarters of a healthcare delivery organisation.

From the above-mentioned description of the intended validators, this study strived to include various SMEs from different backgrounds to establish a diverse pool of validators. This was done to obtain different perspectives for the validation of the HCDMMM. SMEs from different companies or institutions, sectors and countries whose occupation or background is relevant to this study were contacted. The list of SMEs who agreed to participate in the study and their characteristics are included in Table 6.6.

Table 6.5: Validation questionnaire questions

Question no	**Validation dimension**	**Statement or question**
1.1	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level healthcare entities in the public health sector
1.2	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level healthcare entities in the public health sector
1.3	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level healthcare entities in the private health sector
1.4	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level healthcare entities in the private health sector
2	Practicability	The maturity assessment method of the HCDMMM can be put into action to assess the maturity of a healthcare entity's data management
3.1	Usability	Managers of healthcare entities will find it easy to use the HCDMMM for assessing data management maturity
3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment
4.1	Strengths	What, in your view, are the key strengths of the HCDMMM?
4.2	Weaknesses	What, in your view, are the key weaknesses of the HCDMMM?
4.3	Weaknesses	If the HCDMMM was to fail to achieve its stated aim, what do you think would be the reason for this?

After SMEs were contacted, the validation process commenced, which comprised two stages. The first stage entailed the presentation of the HCDMMM to each SME, explaining the research product and the validation process with the aid of the validation questionnaire, and the second entailed the participating SMEs' input and feedback by means of completing the validation questionnaire. Each of these stages is elaborated on below:

Table 6.6: Validation SMEs

SME	Occupation/relevant background	Sector	Country
1	Product manager, software development and maintenance in the public health sector	Service provider to public healthcare	South Africa
2	Health analyst	NPO public healthcare consultant	Zimbabwe
3	Health programme manager	Service provider to public healthcare	South Africa
4	PhD candidate in Health Economics, Health technology and services research	Academia	Netherlands
5	Clinical data and information manager	Private sector hospital	South Africa
6	Associate professor, innovation for inclusive development in healthcare	Academia	South Africa
7	Chief executive officer and director of provincial department of health	Public health sector	South Africa

1. **Presentation:**

   All the participants of the validation process were taken through the HCDMMM. The different sheets of the HCDMMM were explained, including the landing page, the system level overviews, instructions, domain components, assessment sheets and the results sheets. Navigation between sheets was demonstrated and the assessment instructions were explained. The validation process was also explained through the use of the validation questionnaire. After the presentation, a discussion was followed to clarify any remaining questions about anything pertaining to the workings of the HCDMMM and the validation process.

2. **Participant input/feedback:**

   After the presentation, participants were given the opportunity to complete the questionnaire that is illustrated in Appendix I.1. For each question or statement with a scale, the participants had to mark with an ‘x’ how strongly they agreed with the statement or question. Where applicable and/or necessary, free-text space was provided for comments.

The results of the questionnaire that the different SMEs completed are presented in the section that follows. The SMEs' scores are presented and discussed.

6.3.2 Validation results: Responses presentation and discussion

After SMEs gave their feedback, the results were analysed. The results of validation were mainly positive, indicating that the HCDMMM satisfactorily represents the real-world system and that it is useful to the intended user. The results of validation questions 1.1 to 3.2 are given in Table 6.7 and are visually represented in Figure 6.2. The highest score that could be achieved was five and the lowest was one. The average scores presented in the results are also out of five.

The results are presented according to the scores which each SME allocated for the different validation questions and are presented in Table 6.7 and Figure 6.2 per validation question. The average score for each validation question is also illustrated. When SMEs stated that they do not have the experience to score a question it was indicated with a "-" in Table 6.7.

Table 6.7: Validation results

	Validation questions						
	Applicability				Practicability	Usability	
SMEs	**1.1**	**1.2**	**1.3**	**1.4**	**2**	**3.1**	**3.2**
SME 1	4	5	-	-	5	4	4
SME 2	3	3	-	-	5	5	5
SME 3	5	5	5	4	5	4	5
SME 4	5	5	5	5	4	4	4
SME 5	-	-	4	4	3	4	4
SME 6	4	4	4	4	4	4	4
SME 7	4	4	4	4	3	3	3
Question avg	4,17	4,33	4,4	4,2	4,14	4	4,14

The validation results are further discussed during this section. The discussion is structured to present the average allocated score of the validation questions individually and state whether the validation question is satisfactorily addressed. As the SMEs were asked to give their opinion on the strengths and weaknesses of the HCDMMM, and where the HCDMMM might fail with regard to the three validation dimensions, the discussion includes the strengths, weaknesses and where the HCDMMM might fail with regard to each validation question. When a SME allocated a score of three or less it was addressed through the refinements made to the HCDMMM. There was no validation question that was allocated a score of less than three, but SME 2 allocated

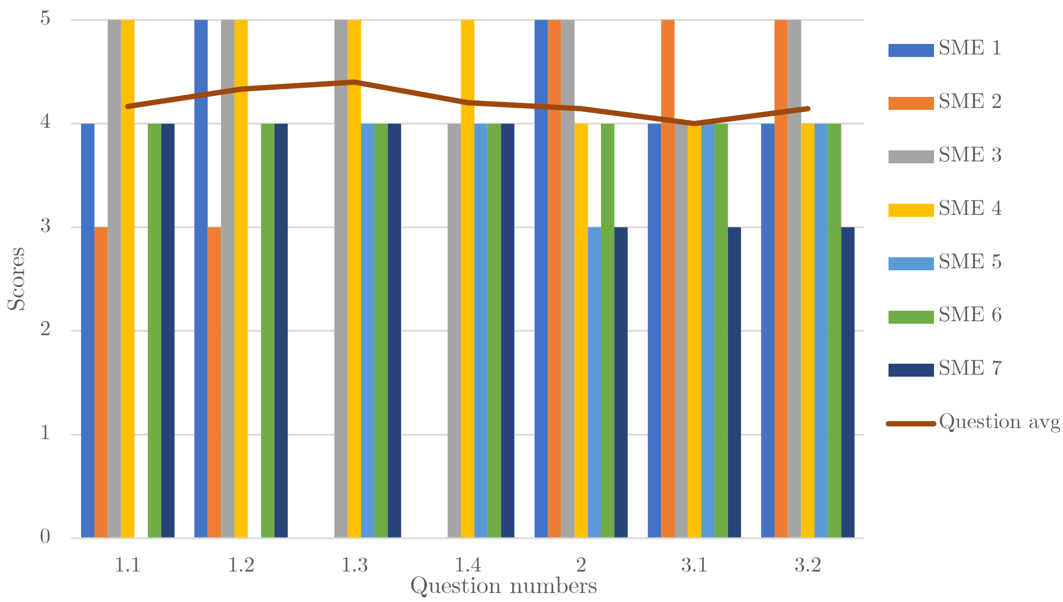


Figure 6.2: Graphical presentation of validation results

validation question 1.1 and 1.2 with a score of three, SME 5 and SME 7 allocated validation question 2 with a score of three, and SME 7 also allocated validation question 3.1 and 3.2 with a score of three. The refinements to improve these scores are discussed in Section 6.3.3. The complete feedback of every validation SME is included in Appendix I in questionnaire form. The rest of this section discusses the applicability dimension (Section 6.3.2.1), followed by the discussion on the practicability dimension (Section 6.3.2.2) and lastly, the usability dimension is discussed (Section 6.3.2.3).

6.3.2.1 Applicability dimension scores, strengths, weaknesses and where the HCDMMM might fail with regard to applicability

Question 1.1 received an average score of 4,17. This validates the HCDMMM is representative of the facility level in the public sector. The average allocated score of question 1.2 was 4,33. This validated the HCDMMM is representative of the organisational level in the public sector. SME 5 stated that she does not have the experience to score question 1.1 and 1.2 as she is employed in the private sector.

A strength that will be advantageous in the public sector that SME 1 pointed out was because the HCDMMM is not limited to only electronic data and considers other data formats. SME 4 believed that other strengths of the HCDMMM that are beneficial in the public sector are that it can be used without access to the internet and that it can be used on low-performance devices. Another strength that SME 4 highlighted with regard to the public sector in general was that the HCDMMM can be used by World Health Organization (WHO) or the Global Fund to assess how facilities in developing countries manage their data. SME 4 believes that trends and patterns will surface across the different facilities, which can be used to determine which capability areas in general they should invest in and improve on.

A point of interest is that SME 2 was neutral about whether the HCDMMM represented the healthcare public sector and gave both question 1.1 and question 1.2 a score of three. A weakness she mentioned was that she felt the HCDMMM is too generic and would add more value if it were tailored more specifically to a public health system and not address private data management components too, as they are quite different and function in different ways. However, the option of "n/a" in the assessment methodology of the HCDMMM provides for these differences between the private and public sector entities and the maturity assessor can determine whether capability areas are relevant to include in the maturity assessment. She also stated that public sector facilities in developing countries capture only country-wide indicators which are used to develop reports and inform decisions on the organisational level. According to her, facilities in the public sector of developing countries do not currently use data they capture. Entities on the organisational level do not collect data, but determine which data facilities collect and pull these data sets for their use. The facilities are at the mercy of the decisions the organisation makes based on the data and they do not make autonomous decisions. The relation between the organisational level and facility level is further addressed in Section 6.3.3 where the refinements to the HCDMMM due to validation are discussed. Another reason SME 2 mentioned why the HCDMMM might fail with regard to validation questions 1.1 and 1.2, was that many data management systems in developing countries, like Zimbabwe, are too far behind for such a tool to be applicable.

Question 1.3 was allocated with an average score of 4,4, not taking into consideration SME 1 and SME 2 who both stated that their experience resides in the public domain and that they cannot give an informed score. Therefore, it is validated that the HCDMMM represents the facility level of the private sector. SME 5 mentioned several points of interest. She stated that the applicability of the HCDMMM depends on the size of the private healthcare facility or group it belongs to. Some of the elements included in the HCDMMM might reside only within the centralised structure of bigger enterprises and are not in the control of the facility. The option of "n/a" in the assessment methodology of the HCDMMM provides for these differences between different entities. She also mentioned that doctors in the private sector work independently and have their own data management systems, which is a key difference from the public sector, and the HCDMMM might be relevant to these individual providers who also work in the facility. However, it falls outside of the scope of this study to validate the applicability of the HCDMMM with regard to the healthcare data management systems of individual providers in private hospitals.

The last applicability question, question 1.4, received an average score of 4,2. SME 1 and 2 were not taken into consideration for this question, as they stated that they do not have the experience to score this question. SME 5 highlighted a point of interest with regard to data capturing and storage for management and decision-making, which is that data is not always collected for the primary purpose of management and decision-making, but data collected in operational processes are reported for management and decision-making. Another aspect she mentioned was that in data storage and sharing there is a contractual component that should be addressed from a POPIA or GDPR perspective. There should also be a consent consideration in place for storage or processing of any demographic or clinical data. Reporting data for management and decision-making as opposed to collecting data for management and decision-making, and the contractual component and consent considerations of data storage and sharing are further addressed in Section 6.3.3 where the refinements to the HCDMMM due to validation are discussed.

With regard to the applicability of the HCDMMM in the private sector in general, SME 5 stated that one of the HCDMMM's strengths is that it covers key areas pertinent to healthcare data. On the other hand, SME 5 stated that a weakness of the HCDMMM is that due to the vast number of systems in healthcare, a high-level view on data management maturity might not provide sufficient direction on where a specific quality gap is and further assessment will be required. Another weakness she mentioned is that the HCDMMM does not take a sociotechnical perspective on data management as people make up a large component to attaining data.

All the applicability questions (question 1.1, 1.2, 1.3 and 1.4) received an average score above four. It is therefore concluded that the HCDMMM is applicable in the public and private health sectors of developing countries, both on the facility and organisation level. Strengths relating to the applicability dimension as a whole include: (i) the HCDMMM captures the main domain components of data management; (ii) the domain components are applicable; (iii) and the HCDMMM is comprehensive and informative. The incorporated refinements with regard to the applicability dimension are outlined in Section 6.3.3.

6.3.2.2 Practicability dimension score, strengths, weaknesses and where the HCDMMM might fail with regard to practicability

The next validation dimension was assessed through question 2 which scored an average of 4,14. This validated the practicability of the HCDMMM and its assessment method, that it can be put into action to assess the maturity of a healthcare entity's data management. SME 3 mentioned that "the model can be used to assess the current state of maturity of both health facilities and organisational levels." A weakness that SME 1 mentioned was that the

HCDMMM requires broad data management knowledge and that in the public sector, with specific reference to rural areas, computers with propriety software are not a given. A reason why SME 3 and 4 said that the HCDMMM might fail with regard to practicability was that in many developing countries, appointed data managers are not always qualified for their role, and might lack the knowledge and skill to use the HCDMMM, whether it is that they do not understand the descriptions for scoring capability areas or interpreting the results. Another concern SME 4 had was that some managers might not be honest when evaluating the capability areas. SME 5 gave this dimension a score of three as she felt it might be difficult to identify the right person to use the tool to make an assessment as this person will require a view across multiple business areas or departments to get an accurate result. SME 1 stated that the success of the HCDMMM is determined by the improvements that are applied to the identified underperforming areas, but in the public sector, healthcare facilities and organisations are understaffed and the suggested improvements might not be followed up on. The refinements regarding the honesty of maturity assessors, their competence and the following up on suggested improvements are discussed further in Section 6.3.3.

6.3.2.3 Usability dimension scores, strengths, weaknesses and where the HCDMMM might fail with regard to Usability

To determine the usability of the HCDMMM, two questions were asked. Question 3.1 received an average score of 4, which validated that managers of healthcare entities will find it easy to use the HCDMMM for assessing data management maturity. SME 1 stated the “model is easy to understand, the instructions are clear and the drop-down arrows will ensure only valid counts are captured.” SME 2 stated that once the HCDMMM is tailored for a specific audience, the tool will be easy to use. She also said the “sheet is well laid out, practical and easy to use.” SME 3 also believed it is usable, giving it a score of 4, but mentioned a weakness that he was concerned it might be time consuming to complete and requires effort. SME 5 commented the tool is easy to understand and complete and SME 6 commented that clarity and simplicity are strengths of the HCDMMM.

Question 3.2 scored an average of 4,14. This validated it is easy to interpret the results of the HCDMMM’s assessment. The strengths that the SMEs pointed out with regard to the results of the HCDMMM was that the summaries of results provide excellent visibility on which areas to focus on (SME 1) and that the radar charts give excellent visualisation of the results (SME 2, SME 3, SME 5 and SME 6). One concern SME 4 and SME 5 mentioned was that due to managers that might be unqualified or who do not have the necessary knowledge or skills, they might have difficulty interpreting the results. SME 5 suggested detail should be added to what the green, yellow and red coloured cells of the overall scores indicate and also what to look out for on the radar graphs. She also suggested replacing the capability category abbreviations in the overall results table headings with their actual names to avoid users having to cross reference back to previous sheets. SME 7 also suggested that using more diagrams can improve the HCDMMM further. How these suggestions were incorporated in the HCDMMM is discussed in Section 6.3.3.

Question 3.1 and 3.2 received an average score of equal to or above four. It is therefore concluded that the HCDMMM is usable. The refinements with regard to usability are outlined in Section 6.3.3.

6.3.3 Refinement of the HCDMMM due to validation

During the validation of the HCDMMM it was validated that the HCDMMM is applicable to the public and private sectors, both on the facility and organisational levels, that its assessment method is practicable and that the HCDMMM is usable. All the SMEs gave predominantly positive feedback, but they also suggested some refinements to improve the HCDMMM further. These refinements are discussed in Table 6.8.

Table 6.8: Refinements due to validation

Refinements	**How the refinements were incorporated**
Applicability	
The inclusion of capability areas related to health indicators under the data collection and data governance domain components	Capability areas related to health indicators were included in the data collection and data governance domain components of the HCDMMM. The capability area to specify the health indicators that the facilities of the organisation should collect was included in the data governance domain component of the organisational level. The capability area to collect the health indicators on the facility level was included in the data collection domain component
The refinement of capability areas related to data collection for management and decision-making in the data collection and data storage domain components	The capability area of data collection for management and decision-making in the data collection domain component on the organisational level was removed from the HCDMMM as SME 2 and SME 5 stated that entities on the organisational level do not collect data for management and decision-making, but that data collected in operational processes is reported for management and decision-making. Therefore, the capability area of storage of data for management and decision-making in the data storage domain component was refined to data aggregation and storage for management and decision-making. The usage of this data is described by the different primary activities in the data usage domain component

Continued on next page

Continued from previous page	
Refinements	**How the refinements were incorporated**
The refinement of maturity level descriptions under data storage and data sharing to include the contractual component and consent consideration for storage or processing of demographic or clinical data	As SME 5 suggested, the contractual component and the consent consideration that should be in place for storage or processing of any demographic or clinical data were included. The contractual component and consent consideration were included in maturity level three of the alignment to data storage and sharing standards, policies and regulations capability areas in the data storage and data sharing domain components
Practicability	
The inclusion of third-party assistance for maturity assessments to ensure the accuracy of assessment results	SME 4 suggested that organisations such as the WHO or Global Fund who are invested in the improvement of healthcare in developing countries can be included to assist the managers with assessments to ensure they make honest and accurate assessments. These organisations are also interested in the assessment results as they will be able to find trends and patterns across facilities and can identify which capability areas should be focused on in general for investment and improvement endeavours. Through the assistance of competent third-parties to assess the healthcare entities' data management, it will be ensured that the results of the HCDMMM are accurate. Including the assistance of third-parties will also ensure accountability that the suggested improvements will be followed up on
Usability	
The improved structuring of the introductory and model overview sheets of the HCDMMM transfer media	To make the introductory and overview sheets of the HCDMMM increasingly user friendly, headings to each section that describe a different component of the HCDMMM were included. The headings give an indication of what each section focuses on. The introductory and overview sheets were also made more concise by removing unnecessary information or removing information that was reiterated. This contributes to prevent reading fatigue and improves the usability of the HCDMMM.
	Continued on next page

Continued from previous page

Refinements	**How the refinements were incorporated**
The inclusion of results indicators to assist HCDMMM users in interpreting the data and identifying the components to focus on for improvement	To assist HCDMMM users in interpreting the assessment results and identifying the components to focus on for improvement, result indicators were included in the Results sheets. The result scores are highlighted according to a green-yellow-red colour scale in the results tables which indicates: (i) which capability areas scored the best; (ii) which capability areas received average scores; and (iii) which capability areas scored the lowest scores. Example radar charts that give an indication of how to interpret the results radar charts were also included. Charts were included that indicate: (i) excellent results; (ii) good results; (iii) results that cause reasonable concern; and (iv) results that cause great concern. The headings of the overview of results table were also changed from using the abbreviations of primary activities, enabling practices and supporting structures to writing them out in full with their abbreviations

6.4 Conclusion on the verification and validation chapter

The verification process determined whether the HCDMMM was theoretically sound and that it adhered to the requirements specifications. Theoretical verification was conducted in two phases. The first phase was executed during the iterative development process where the focus was to verify developed components as they were completed. The second phase was conducted after the HCDMMM was completed where the focus was to establish that the HCDMMM as a whole was theoretically sound. It was concluded that the HCDMMM is theoretically sound and that most requirements were satisfactorily addressed, but not all, which resulted in refinements.

The second phase of verification was establishing whether the HCDMMM addressed the requirements specifications. It was concluded that all the requirements were addressed satisfactorily.

It was validated whether the HCDMMM represents the real-world system and that it is useful to the intended users. The three dimensions the HCDMMM was validated according to included applicability, practicability and usability. It was concluded that the HCDMMM is applicable to the public and private health sectors of developing countries, both on the facility and organisation level, that it is practicable and usable, although final refinements were included to improve the HCDMMM further. The final HCDMMM is presented in Chapter 5.

Chapter 7

Conclusion

Chapter 7 consists of an overview of the research done during this study (Section 7.1), an evaluation of the research objectives (Section 7.2), and the limitations of the study (Section 7.3). It also discusses possible opportunities for future research in this field of study (Section 7.4). Lastly, this chapter is concluded (Section 7.5).

7.1 Overview of research

The goal of this research study is to contribute towards the improvement of healthcare data management in developing countries in order to improve healthcare delivery. The outcome of this study was a proposed maturity model for healthcare delivery entities on the facility level such as hospitals and clinics and their headquarters on the organisational level. The Healthcare Data Management Maturity Model (HCDMMM) assists the specific healthcare entities assess their data management.

In Chapter 1 the data management challenges that developing countries face were introduced. The research problem was stated, as well as the research objectives. The research strategy was also described. Chapters 2, 3 and 4 comprised the literature review which contributed to the development of the HCDMMM as described in Chapter 5. In Chapter 6 the evaluation of the HCDMMM is described.

Chapter 2 focused on the exposition of healthcare data management. Firstly, the healthcare system was described, followed by an explanation of data management in healthcare. The healthcare data management challenges were also scoped to determine all the challenges that impede the effective management of healthcare data in developing countries.

Chapter 3 included the development of the requirements specification that a possible research product should address. To do this it was necessary to determine the challenges landscape to illustrate all the challenges of healthcare data management across the value chain. Therefore, the healthcare data value chain was constructed based on literature on big data value chains. The developed healthcare data value chain was used to plot the healthcare data management scope of challenges in the different value chain components. From the challenges landscape and Chapter 2, it was possible to specify the requirements that the research product should address to address the healthcare data management challenges in developing countries.

In Chapter 4 the concept of maturity models was investigated to demonstrate the usefulness of developing a maturity model to assist in addressing the challenges landscape that was constructed in Chapter 3, by assisting healthcare delivery entities assess their data management. The chapter is structured so that it firstly describes the origin, purpose and value of maturity models in general, followed by the basic structure of maturity models. The importance of using a defined methodology with design decisions and design principles to develop maturity models, is then explained. Lastly, maturity models in the healthcare data management domain are then described to learn what has been done in the past, what challenges they strived to address and how the model to be developed in this study can be unique from previous studies.

Chapter 5 described how the HCDMMM was developed to assist in addressing the healthcare data management challenges in developing countries. Firstly, an overview of the development methodology was given. This described the development process as: (i) identifying the need for new opportunity; (ii) defining the scope of the model; and (iii) the design and development of the HCDMMM. This was an iterative development process which consisted of various steps that included a verification step that was described in detail in Chapter 6. After the structure and the content of the HCDMMM was completed and theoretically verified, the transfer media, being the form in which the HCDMMM will be used, was developed to improve its usability and practicability. The future updates of the HCDMMM was also considered and it was evaluated as described in Chapter 6.

Chapter 6 consisted of the evaluation of the HCDMMM. Evaluation was comprised of verification and validation. Verification consisted of two parts, theoretical verification and the requirements specification. Theoretical verification was executed in two steps, the first step was verifying the model structure and content as it was being developed through semi-structured interviews with various relevant knowledge domain SMEs. They verified that the structure and content were theoretically sound and identified components which were lacking or missing. The second step was executed after the development process. This was carried out through a questionnaire that two healthcare data management SMEs completed to verify the HCDMMM comprehensively. The questions were constructed to verify the structure and content of the HCDMMM and that it addressed the specified requirements. This led to a few final refinements, whereafter the transfer media was also developed. The requirements specification was then reconsidered to establish that the HCDMMM addressed all the specified requirements. The second part of the evaluation was validation. The aim of validation was to determine whether the HCDMMM addressed the needs of the user from the users' perspective. This was done through a questionnaire that was constructed to validate the applicability, practicability, usability. Several SMEs from various health sectors completed the questionnaire to validate the HCDMMM from different perspectives.

Lastly, this chapter gives an overview of the research, demonstrates how the research objectives were achieved, addresses the limitations of this study, and suggests possible future work and recommendations.

7.2 Achievement of research objectives

It was envisioned that by addressing the following objectives, the aim of the study would be addressed. In Chapter 1, the research objectives were stated. In Table 7.1, each of the research objectives are stated and an account of the sections in which the research objectives are attained are discussed.

Table 7.1: Achievement of the research objectives

Research objectives	**Account of how the objectives are attained**
1. Describe the context of healthcare and health care data management in order to gain a better understanding of healthcare data management in developing countries	
1.1 Describe the delivery of healthcare as a system	This research objective was achieved in Section 2.1 where the systems approach was first described and then the systems approach with regard to healthcare. Healthcare was described as a system with various elements that carries out different functions on different levels
1.2 Define data management and describe data management in the context of the healthcare sector	In Section 2.2 this objective was achieved. Section 2.2.1 gave a common understanding of what is meant by data management for this study and Section 2.2.2 described the importance of having a data strategy when managing data. Data management in the healthcare sector was briefly described in 2.2.3. Traditional data management and big data in healthcare were described to illustrate the usage of data in the healthcare sector. Data management on different system levels was also described
	Continued on next page

Continued from previous page

Research objectives	**Account of how the objectives are attained**
1.4 Identify the significant challenges of healthcare data management in developing countries	This research objective was achieved in Section 2.3 when the healthcare data management challenges were scoped through a structured literature review. This review yielded the scope of data management challenges in developing countries.
2. Specify the requirements that the proposed research product should address in order to ensure the research product adequately addresses the problem	
2.1 Identify and describe the significant healthcare data management components	To accomplish this research objective, the significant healthcare data management components were identified through reviewing different big data value chains in Section 3.1. By using the different big data value chains, the different relevant data management components for this study were identified and described
2.2 Determine the challenges landscape of data management across the whole healthcare data management value chain	By using the scope of challenges as described in Section 2.3, and the data value chain that was developed, as described in Section 3.1, the challenges landscape was determined in Section 3.1.2 to fulfil this research objective. The challenges landscape was needed in order to specify the requirements that the proposed research product should adhere to. The whole challenges landscape can be seen in Appendix B
2.3 Specify the requirements that the proposed research product should address to be able to assist in identifying healthcare data management components to improve on and to address the problem statement of this study	This research objective was attained through the specification of the requirements in Section 3.2. These requirements were specified through the inputs from literature that are included in Chapters 2 and 3. The challenges landscape determined in Section 3.1.2 was also used as an input to determine the requirements

Continued on next page

Continued from previous page

Research objectives	**Account of how the objectives are attained**
3. Identify, select and describe a suitable research product that is able to facilitate the identification of healthcare data management components to improve on in order to address the healthcare data management problem stated for this study	This objective was achieved in Chapter 4 where a maturity model was identified, selected and described as an appropriate research product. In Section 4.1 the origin, purpose and value of maturity models were described. The basic constructs of maturity models were described in Section 4.2. Why it is important to use a defined methodology with design decisions and principles to develop a maturity model was explained in Section 4.3. The development methodologies as described by De Bruin *et al.* (2005) and Becker *et al.* (2009) was conveyed, along with the design decisions of Mettler (2010*b*) and the design principles of Pöppelbuß and Röglinger (2011). An exposition of past maturity models was given in Section 4.4 to demonstrate how maturity models were used to address healthcare data management challenges. From this exposition it was possible to determine what types of maturity models were developed, which domain areas they focused on, what could be learnt from these models and what gap the proposed research product can address
4. Develop the research product in order to provide an appropriate means to identify the healthcare data management components to focus on for improvement endeavours	
4.1 Use a defined design and development methodology with the appropriate design decisions and principles to develop the research product	The maturity model development methodologies, design decisions and principles described in Section 4.3 were incorporated to develop a maturity model to meet the specified requirements in Section 3.2. The development methodology for this study was described in Section 5.1. The different steps of the development process were then described, which consisted of the identification of the need for a maturity model (Section 5.2), the definition of the scope of the model (Section 5.3), the iterative design and population of the model (Section 5.4), and the description of the final model (Section 5.5)

Continued on next page

Continued from previous page

Research objectives	**Account of how the objectives are attained**
4.2 Verify the research product theoretically and whether it met the specified requirements	To achieve this research objective, the HCDMMM was verified theoretically in Section 6.2.1. Theoretical verification was carried out both during the development of the model structure and content and after its development. This ensured that the structure and the content of the model was theoretically sound. After this the transfer media, as the model will be used by users, was developed and evaluated against the specified requirements to determine whether all the requirements were addressed (Section 6.2.2)
4.3 Validate the applicability, practicability and usability of the developed research product and determine its strengths and weaknesses	This research objective was fulfilled in Section 6.3. Validation involved the use of SMEs to evaluate the HCDMMM against the dimensions of applicability, practicability and usability. This was done using a questionnaire to evaluate the HCDMMM. SMEs also gave their opinions on the strengths and weaknesses of the HCDMMM

7.3 Limitations

Literature was used as a theoretical foundation for the development of the HCDMMM and therefore, the HCDMMM was developed based on secondary information. By using a literature review, the study is limited to merely summarising and organising existing information. Additionally, this study is subject to bias as it employed qualitative methods to verify and validate the HCDMMM through the use of SME inputs from various knowledge domains related to healthcare.

The study focused on the technical components of healthcare data management in developing countries from a strategic perspective and did not include the human components (i.e. equipping staff and senior decision-makers, managing user resistance to new data management practices, data literacy of healthcare workers, addressing cultural barriers and establishing a data driven culture amongst healthcare workers) in the HCDMMM development. This study also did not consider costs and finances with regard to healthcare data management in the development of the HCDMMM.

A limitation with regard to the findings of the study is that it was not evaluated through a real-world case study. A real-world case study would have been able to determine whether managers of healthcare entities are qualified and have the knowledge to use the HCDMMM to make a maturity assessment and to determine whether the healthcare entity has the necessary technology available to make a maturity assessment. As a real-world case study was not conducted, this study is limited to not being able to determine this.

7.4 Future work

For future work the HCDMMM can be expanded to include the human components in the HCDMMM and not only focus on the technical system. Human input plays a significant role in the process of attaining healthcare data and can be included so that the HCDMMM will also be able to assess the maturity of the human components of healthcare data management. Human components like equipping staff and senior decision-makers, managing with user resistance to new data management practices, data literacy of healthcare workers, addressing cultural barriers and establishing a data driven culture amongst healthcare workers can be included.

The HCDMMM can also be applied in a real-world case study to determine its applicability, practicability and usability in a real-world scenario. This will be very beneficial to determine whether real users are able to use the HCDMMM as managers in developing countries might not have the knowledge to use the HCDMMM to make a maturity assessment and develop improvement measures from it. Case studies can include various facility level healthcare entities such as hospitals and clinics in rural or urban areas in developing countries to determine whether the HCDMMM can be used in such settings. Real-world case studies will be able to determine whether managers of healthcare entities are qualified and have the knowledge to use the HCDMMM to make a maturity assessment and to determine whether the healthcare entity has the necessary technology available to make a maturity assessment. Case studies should also include the application of the HCDMMM on the organisational level of a country's data management system.

Future work can also include determining transferability. The degree to which the HCDMMM can be transferred to other contexts or settings can be determined. The transferability of the HCDMMM can be determined through multiple case studies in different developing countries to establish whether the HCDMMM can be transferred to multiple developing country settings.

7.5 Chapter conclusion

This chapter gave a summary of the research involved for this study. It also discussed how the research objectives were reached. The limitations of this study were also discussed, followed by a section on opportunities for future research in this field of study.

Appendix A

Health care data management scope of challenges in developing countries

This Appendix shows an extract of the scope of health care data management challenges that were developed in Microsoft Excel. This extract gives the content of only the integration challenges category. All the other categories follow the same format.

Table A.1: Scope of data integration challenges extract

					10	9	16	11	15	24	6	10	14	11	2	13	2	7	2	5	2	3	0
					162																		
Title	Author	Year	Article challenges	Country	Integration Challenges																		
			211		General Challenges	Multiple not linked patient Records	Lack of system Standardisation	Lack of interoperability	Variety of non integrated (heterogeneous) systems	Fragmented systems	Data organisation challenges	System Complexity	Lack of data aggregation	Lack of centralised data storage	Lack of data warehouse	Paper based system	Data synchronisation	System unsustainability	System design	Flexibility	Decentralisation	Incompatible infrastructure	
Health Management for Sustainable Development	Reshetnikov, A., Fedorova, J., Prisyazhnaya, N., (...), Shamshurina, N., Kolesnichenko, O.	2019	3	Russia								x						x					
Real-Time ECG Monitoring over Multi-Tiered Telemedicine Environment using Firebase	Mukhopadhyay, A., Xavier, B., Sreekumar, S., Suraj, M.	2018	18	India																			
Factors Affecting the Successful Adoption of e-Health Cloud Based Health System from Healthcare Consumers' Perspective	Idoga, P.E., Toycan, M., Nadiri, H., Çelebi, E.	2018	15	Nigeria					x														
Private cloud solution for Securing and Managing Patient Data in Rural Healthcare System	Ganiga, R., Pai, R.M., Pai, M., Sinha, R.K.	2018	25	India						x						x							
Investigating TOE Factors Affecting the Adoption of a Cloud-Based EMR System in the Free-State, South Africa	Masana, N. & Muriithi, G.M.	2018	8	South Africa						x		x		x		x							
Understanding health worker data use in a South African antiretroviral therapy register	Jobson, G., Murphy, J., van Huyssteen, M., Myburgh, H., Hurter, T., Grobbelaar, C.J., Struthers, H.E., McIntyre, J.A. & Peters, R.P.H.	2018	10	South Africa																			
Feasibility Analysis for Deploying National Healthcare Information System (NHIS) for Pakistan	Yaqoob, T., Mir, F., Abbas, H., (...), Shafqat, N., Amjad, M.F.	2017	29	Pakistan	x		x	x		x		x											
Lessons from implementation of ecohealth projects in Southern Africa: A principal investigator's perspective	Chimbari, M.J.	2017	15	Southern Africa			x			x	x												

Quality of routine health data collected by health workers using smartphone at primary health care in Ethiopia	Medhanyie, A.A., Spigt, M., Yebyo, H., (...), Dinant, G.-J., Blanco, R.	2017	16	Ethiopia			x																
Health data integration with Secured Record Linkage: A practical solution for Bangladesh and other developing countries	Khan, S.I., Hoque, A.S.M.L.	2017	20	Bangladesh	x	x			x	x													
An Analysis of the Problems for Health Data Integration in Bangladesh	Khan, S.I., Hoque, A.S.Md.L.	2017	11	Bangladesh	x	x	x	x	x	x													
Towards sustainable mHealth applications for maternal and child health: The case of Sahabat Bundaku-an integrated mobile application for mothers and midwives	Irawan, Y.S., Soegijoko, S., Koesoema, A.P., (...), Isdiningrat, I.S., Husin, F.	2017	15	Indonesia	x*																		
Challenges for using IT in Mexico's health care industry (Aguascalientes México case)	Vivanco, J.S., González, M., Oropeza, M.A.	2017	5	Mexico						x													
INTEGRATE-HTA: A low- and middle-income country perspective	Bijlmakers, L., Mueller, D., Kahveci, R., Chen, Y., Van Der Wilt, G.J.	2017	7	LMIC																			
Patchworks of logistics management information systems: Challenges or solutions for developing countries?	Bergum, B.-I., Nielsen, P., Sæbø, J.I.	2017	12	Uganda, Tanzania			x	x	x	x			x			x							
Privacy implications in wearable healthcare monitoring devices	Moganedi, S., Dlamini, Z.	2017	12	South Africa										x									
Clinical characteristics and outcomes of patients with stroke admitted in tertiary hospitals in Zimbabwe: A retrospective one-year study	Kaseke, F., Stewart, A., Gwanzura, L., Hakim, J., Chikwasha, V.	2017	9	Zimbabwe	x											x							
Diabetic data analysis in healthcare using Hadoop architecture over big data	Sampath, P., Tamilselvi, S., Saravana Kumar, N.M., Lavanya, S., Eswari, T.	2017	4	India							x												
Factors impacting the adoption decision of health data standards in tertiary healthcare organisations in Saudi Arabia	Alkraiji, A.I., Jackson, T., Murray, I.	2016	15	Saudi Arabia	x		x	x		x	x	x											

Adapting HIV patient and program monitoring tools for chronic non-communicable diseases in Ethiopia	Letebo, M., Shiferaw, F.	2016	9	Ethiopia		x					x		x										
Integration of smart wearable mobile devices and cloud computing in South African healthcare	Mvelase, P., Dlamini, Z., Dludla, A., Sithole, H.	2016	23	South Africa						x	x*			x*									
Diabetes in Thailand: Status and Policy	Reutrakul, S., Deerochanawong, C.	2016	1	Thailand																			
Privacy and security problems of national health data warehouse: A convenient solution for developing countries	Khan, S.I., Hoque, A.S.M.L.	2016	21	Bangladesh		x	x		x	x					x								
Exploring the information and ICT Skills of health professionals in low- and middle-income countries	Koivu, A., Mavengere, N., Ruohonen, M.J., Hederman, L., Grimson, J.	2016	18	LMIC																			
Big data in third world countries: Do the means justify the end?	Gupta, V.G., Bora, G.S., Mavuduru, R.S.	2016	7	India										x									
A blank page: Feedback from first referral hospitals to primary health care clinics	Legodi, T.L., Wolvaardt, J.E.	2015	3	South Africa																			
Implementation of a cloud-based electronic medical record to reduce gaps in the HIV treatment continuum in rural Kenya	Haskew, J., Rø, G., Turner, K., (...), Sirengo, M., Sharif, S.	2015	20	Kenya			x*	x*		x													
Success criteria for electronic medical record implementations in low-resource settings: A systematic review	Fritz, F., Tilahun, B., Dugas, M.	2015	31	African countries	x		x	x		x	x		x					x					
Implementation of Policies and Strategies for Control of Noncommunicable Diseases in Malawi: Challenges and Opportunities	Lupafya, P.C., Mwagomba, B.L.M., Hosig, K., Maseko, L.M., Chimbali, H.	2015	2	Malawi	x					x													
Development and evaluation of an integrated electronic data management system in a South African metropolitan critical care service	Allorto, N.L., Wise, R.D.	2015	23	South Africa																			
Practicing evidence-based medicine in developing countries	Puustjärvi, J., Puustjärvi, L.	2015	1	Developing countries																			

Evaluation of antiretroviral treatment programme monitoring in Eastern Cape, South Africa	Kaposhi, B.M., Mqoqi, N., Schopflocher, D.	2015	22	South Africa									x										
A structured approach to recording AIDS-defining illnesses in Kenya: A SNOMED CT based solution	Oluoch, T., de Keizer, N., Langat, P., (...), Kwaro, D., Cornet, R.	2015	7	Kenya			x																
Role of OpenEHR as an open source solution for the regional modelling of patient data in obstetrics	Pahl, C., Zare, M., Nilashi, M., (...), Supriyanto, E., Ibrahim, O.	2015	6	Brazil				x															
Investigating factors for Health Information Systems use by physicians in developing countries	Adenuga, O.A. & Kekwaletswe, R.M.	2015	15	African countries																			
Cloud computing impact on heathcare in South Africa (Cloud Computing: Impact on CIOs and the IT Team)	Mgozi, T., Weeks, R.	2015	20	South Africa			x	x	x	x							x	x					
Distributed data managing in health care social network based on mobile P2P	Wang, Y., Liu, H., Wang, L.	2015	7	China								x								x	x		
Creative intelligence: Does intelligence in health care always have to mean high-tech?	Fischer, S.	2014	4	Developing countries																			
Using mobile computing to support malnutrition management in South Africa	Nyumbeka, D., Wesson, J.	2014	16	South Africa																			
Efficacy of knowledge transfer and exchange between CHWs and community members in low resource settings	Warkulwiz, G., Paul, L., Mehta, K.	2014	22	Kenya																			
Supporting research sites in resource-limited settings: Challenges in implementing information technology infrastructure	Whalen, C.J., Donnell, D., Tartakovsky, M.	2014	20	Resource Limited Countries										x			x						
Critical information technology issues in Turkish healthcare	Turan, A.H., Palvia, P.C.	2014	34	Turkey		x		x	x	x		x				x						x	
Infrastructuring work: Building a state-wide hospital information infrastructure in India	Aanestad, M., Jolliffe, B., Mukherjee, A., Sahay, S.	2014	29	India		x*	x		x	x		x	x	x		x		x	x	x			
Globalization and ICTs: Potentials and challenges for the public health sector of developing countries	Mengiste, S.A.	2013	24	Ethiopia			x			x			x			x		x		x	x		

E-health implementation challenges in Iranian medical centers: A qualitative study in Iran	Sharifi, M., Ayat, M., Jahanbakhsh, M., (...), Mokhtari, H., Wan Ismail, W.K.	2013	28	Iran			x			x		x											
Information technology systems in public sector health facilities in developing countries: The case of South Africa	Cline, G.B., Luiz, J.M.	2013	22	South Africa					x			x		x		x							
Human factors affecting the quality of routinely collected data in South Africa	Nicol, E., Bradshaw, D., Phillips, T., Dudley, L.	2013	18	South Africa									x			x							
Cloud-enabled search for disparate healthcare data: A case study	Bhaskaran, S., Suryanarayana, G., Basu, A., Joseph, R.	2013	15	India				x	x	x			x										
Combating rural child malnutrition through inexpensive mobile phones	Medhi, I., Jain, M., Tewari, A., (...), Matheke-Fischer, M., Cutrell, E.	2012	32	India		x							x	x									
Implementing an electronic health record as an objective measure of care provider accountability for a resource-poor rural area in the Dominican Republic	Palmer, R.C., Simms-Cendan, J.S.	2012	16	Dominican Republic						x			x	x		x							
CRHIS: Cloud based rural healthcare information system	Patra, M.R., Das, R.K., Padhy, R.P.	2012	27	India		x				x						x				x		x	
A standards-based open source application to gather health assessment data in developing countries	Gainer, A., Roth, M., Strong, P., Davis, J.	2012	19	Ethiopia			x		x				x							x			
Improving quality and use of data through data-use workshops: Zanzibar, united republic of Tanzania	Braa, J., Heywood, A., Sahay, S.	2012	16	Tanzania			x		x	x		x			x			x					
Data integrity module for data quality assurance within an e-health system in sub-Saharan Africa	Monda, J., Keipeer, J., Were, M.C.	2012	16	Sub-Saharan Africa				x		x													
An IS contribution to the UN Millennium Development Goals: Next generation vaccination management in the developing world	Wang, W., Elliot, S., Williams, M.-A.	2010	34	Developing countries	x																	x	
Modeling access to healthcare: Design research in remote communities	Bindle, S., Joost, G.	2010	11	India																			

A web-based electronic health care system for the treatment of pediatric HIV	Paul, S., Das Bhattacharya, S., Sudar, A., (...), Mukhopadhyay, J., Majumdar, B.	2009	10	India									x	x		x							
Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa	Mate, K.S., Bennett, B., Mphatswe, W., Barker, P., Rollins, N.	2009	22	South Africa					x				x						x				
PRISM framework: A paradigm shift for designing, strengthening and evaluating routine health information systems	Aqil, A., Lippeveld, T., Hozumi, D.	2009	24	Developing countries					x									x					
Romanian health system strategic directions for the next decade	Vlădescu, C., Galan, A., Olsavszky, V., Scîntee, S.G.	2009	4	Romania	x				x														
Towards a knowledge management framework for assisting organisations to evaluate their own (non-clinical) approaches to the dissemination of knowledge about HIV/AIDS intervention programmes in South Africa	Sassman, R., Lehaney, B., Marshall, I.	2008	11	South Africa									x			x							
Perspectives on high technologies for low-cost healthcare: The Chinese scenario	Poon, C., Zhang, Y.-T.	2008	3	China		x								x									
				Challenges per year																			Total category challenges per yr
				2008	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	4
				2009	1	0	0	0	3	0	0	0	2	1	0	1	0	1	1	0	0	0	10
				2010	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2
				2011	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
				2012	0	2	2	1	2	4	0	1	3	2	1	2	0	1	0	2	0	1	24
				2013	0	0	2	1	2	3	0	2	3	1	0	3	0	1	0	1	1	0	20
				2014	0	2	1	1	2	2	0	2	1	2	0	2	1	1	1	1	0	1	20
				2015	2	0	4	4	1	4	1	1	2	0	0	0	1	2	0	1	1	0	24
				2016	1	2	2	1	1	3	3	1	1	2	1	0	0	0	0	0	0	0	18
				2017	5	2	5	3	3	6	2	1	1	1	0	2	0	0	0	0	0	0	31
				2018	0	0	0	0	1	2	0	1	0	1	0	2	0	0	0	0	0	0	7
				2019	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	2

Appendix B

Challenges landscape

This appendix indicates to what parts of the data value chain all the different challenges that were identified in the scope of challenges belongs. This forms the complete challenge landscape.

Table B.1: Challenge landscape

Identification	Challenge	Data collection	Data sources	Data acquisition	Data transmission	Data curation	Data storage	Data analysis	Data usage	Master data management	Data life cycle management	Data security and privacy management	Technology and infrastructure	Finances and costs	Human contributions	Social(S)/Technical (T) factor
V1	No collection	x													x	S;T
V2	Collection delay	x													x	S;T
V3	Errors with collection	x								x					x	S;T
V4	Inefficient collection processes and methods	x								x						S;T
V5	Lack of standardised/ proper entry forms	x								x						T
V6	Data collected multiple times	x								x						T
V7	Variety in type of data between health care facilities	x	x		x	x				x						T
V8	No digital data capturing	x													x	T
V9	Copying from paper to digital errors	x													x	S
V10	Time constraint with collection	x													x	S;T
V11	Human regard for quality collection	x													x	S
V12	Skill required with collection	x													x	S

Continued on next page

Continued from previous page

ID	Challenge	DC	DS	DT	DA	DC	DS	DA	DU	MDM	DLCM	DSPM	T/I	F/C	HC	S/T
V13	Lack of data capturers	x													x	S
V14	Lack of collection equipment	x											x			T
V15	Delay in digitalising paper based data	x											x		x	S;T
V16	Not time effective collection	x											x		x	S;T
V17	Lack of supervision and verification of data collection	x								x					x	S
V18	Collection security	x										x				S;T
V19	Inability to tracking/monitoring	x								x			x			T
V20	Lack of standardised monitoring tools	x								x			x			T
V21	Paper based monitoring	x											x			T
V22	Monitoring is time consuming	x											x		x	S;T
V23	Lack of real time monitoring	x	x	x	x								x			T
V24	Lack of remote monitoring	x	x	x	x								x			T
V25	Bad quality of data for monitoring	x				x				x						T
V26	Need for up-to-date data	x									x					T
V27	Required monitoring skills	x													x	S
V28	Historical data inaccessibility		x				x						x			T
V29	Paper based systems		x						x				x			T
V30	Real time data inaccessibility		x	x			x						x			T
V31	Lack of remote data access		x	x			x		x				x			T

Continued on next page

Continued from previous page

ID	Challenge	DC	DS	DT	DA	DC	DS	DA	DU	MDM	DLCM	DSPM	T/I	F/C	HC	S/T
V32	Difficulty of accessibility		x	x			x		x				x			T
V33	Inefficient accessibility		x	x			x						x			T
V34	Ineffective accessibility		x	x			x						x			T
V35	Lack of timely accessibility		x	x			x			x			x			T
V36	Unauthorised access		x				x		x			x				S;T
V37	Data not shared with different users				x		x		x				x			T
V38	Data not disseminated to patients				x				x				x			T
V39	Transmission latency				x								x			T
V40	Network unreliability				x								x			T
V41	Network insecurity				x							x	x			T
V42	Software issues				x								x			T
V43	Transmission errors				x	x										S;T
V44	Form of transmission				x					x						T
V45	Unintelligible data transmission				x	x				x						T
V46	Inaccurate information				x	x				x						S;T
V47	Data sharing security issues				x							x				S;T
V48	Paper based transmission				x								x			T
V49	Data not submitted				x										x	S
V50	Multiple not linked patient records		x			x				x						T
V51	Lack of system standardisation		x							x						T
V52	Lack of interoperability					x				x						T
V53	Variety of non integrated (heterogeneous) systems		x			x				x						T
V54	Fragmented systems		x			x				x						T

Continued on next page

Continued from previous page

ID	Challenge	DC	DS	DT	DA	DC	DS	DA	DU	MDM	DLCM	DSPM	T/I	F/C	HC	S/T
V55	Data organisation challenges					x				x						T
V56	System Complexity					x				x						T
V57	Lack of data aggregation			x	x	x				x						T
V58	Lack of centralised data storage					x	x				x					T
V59	Lack of data warehouse					x	x				x					T
V60	Paper based system		x										x			T
V61	System unsustainability	x	x	x	x	x	x	x	x	x	x	x	x	x	x	S;T
V62	System design	x	x	x	x	x	x	x	x	x	x	x	x	x	x	T
V63	Flexibility		x			x	x	x	x	x			x			T
V64	Decentralisation		x	x	x	x	x			x			x			T
V65	Data synchronisation					x	x						x			T
V66	Incompatible infrastructure			x		x							x			T
V67	Bad process methods					x				x					x	S;T
V68	Lack of required processing tools					x							x			T
V69	Timeliness of processing					x							x		x	S;T
V70	Poor Processing					x							x		x	S;T
V71	Lack of storage Backups						x						x			T
V72	Lack of digital data storage						x						x			T
V73	Remote storage						x						x			T
V74	Storage infrastructure						x						x			T
V75	Database scalability						x				x		x			T
V76	Stored data out of date						x				x					T
V77	Loss of stored data						x				x					T
V78	Reliability of storage						x			x	x					T

Continued on next page

Continued from previous page

ID	Challenge	DC	DS	DT	DA	DC	DS	DA	DU	MDM	DLCM	DSPM	T/I	F/C	HC	S/T
V79	Paper based and electronic databases (redundant)						x			x						T
V80	Poor storage techniques						x			x						S;T
V81	Unstructured storage		x			x	x			x						T
V82	Size of data					x	x			x			x			T
V83	Storage delay				x		x									S;T
V84	Storage security						x					x				S;T
V85	Issues with analysis data					x		x		x						T
V86	Unreliable analysis					x		x		x					x	S;T
V87	Bad quality of analysis					x		x		x					x	S;T
V88	Inefficient analysis							x		x					x	S;T
V89	Errors in analysis							x					x		x	S;T
V90	Lack of timely analysis							x					x		x	S;T
V91	Bad analysis methods							x		x			x		x	S;T
V92	Analysis tools needed							x					x			T
V93	Paper based analysis							x					x			T
V94	Required analysis skills							x							x	S
V95	Reporting skills required								x						x	S
V96	Paper based reporting								x				x			T
V97	Misreporting								x				x		x	S;T
V98	Inefficient reproting								x	x			x		x	S;T
V99	Needed functionalities for reporting								x	x			x			T
V100	Procuring reporting registers difficulty				x				x	x			x			T
V101	Unreliable reporting								x	x						S;T
V102	Lack of timely reporting								x	x			x		x	S;T
V103	Lack of standard reporting format								x	x						T
V104	Incomplete data for reporting	x	x			x	x		x							S;T

Continued on next page

Continued from previous page

ID	Challenge	DC	DS	DT	DA	DC	DS	DA	DU	MDM	DLCM	DSPM	T/I	F/C	HC	S/T
V105	Out of date reports/ registers								x		x					T
V106	Usability					x			x	x						T
V107	Unused data								x	x					x	S;T
V108	Required utilization skills								x						x	S
A1	Legislation and regulations									x						T
A2	Policies and standards									x						T
A3	Leadership									x					x	S
A4	Lack of innovation adoption									x					x	S;T
A5	Collaboration									x					x	S;T
A6	Missing frameworks									x						T
A7	Supervision									x					x	S;T
A8	Quality standards					x				x						T
A9	Missing/partial data	x								x						S;T
A10	Inaccurate data	x								x						S;T
A11	Data duplication	x				x										S;T
A12	Unstructured data					x										S;T
A13	Data noise			x		x				x						T
A14	Unintelligible data			x		x				x						T
A15	Unreliable data			x		x										S;T
A16	Required human skills	x				x				x					x	S
A17	Data discrepancy/ inconsistency	x				x				x						S;T
A18	Supervision	x				x				x					x	S
A19	Data timeliness	x			x	x				x					x	S;T
A20	Data relevance	x		x		x				x					x	S;T
A21	Patient privacy	x	x				x					x			x	S;T
A22	Data integrity	x	x	x	x	x	x					x			x	S;T
A23	Malicious behaviour	x	x	x	x	x	x					x			x	S
A24	Accountability									x		x			x	S;T
A25	Required security skills											x			x	S
A26	Anonymity	x	x				x					x				T
A27	Authorization						x					x				T

Continued on next page

Continued from previous page

ID	Challenge	DC	DS	DT	DA	DC	DS	DA	DU	MDM	DLCM	DSPM	T/I	F/C	HC	S/T
A28	Confidentiality	x	x				x					x				T
A29	Security technology needed	x	x				x					x	x			T
A30	Differential access		x				x		x			x	x			T
A31	Malware		x				x					x	x			T
A32	Security standards									x		x				T
A33	Policies									x		x				T
A34	Legislation									x		x				T
A35	Network unavailability	x	x	x	x	x	x	x	x				x			T
A36	Lack of power supply	x	x	x	x	x	x	x	x				x			T
A37	Infrastructure	x	x	x	x	x	x	x	x				x			T
A38	Computers	x	x	x	x	x	x	x	x				x			T
A39	Hardware availability	x	x	x	x	x	x	x	x				x			T
A40	Software issues	x	x	x	x	x	x	x	x				x			T
A41	Required technological skills	x	x	x	x	x	x	x	x				x			S
A42	Robustness	x	x	x	x	x	x	x	x				x			T
A43	Compatability	x	x	x	x	x	x	x	x	x			x			T
A44	Supplimentary tools	x	x	x	x	x	x	x	x	x			x			T
A45	Power cost	x	x	x	x	x	x	x	x					x		T
A46	Infrastructure cost	x	x	x	x	x	x	x	x					x		T
A47	Network cost	x	x	x	x	x	x	x	x					x		T
A48	Technology cost	x	x	x	x	x	x	x	x					x		T
A49	Maintenance cost	x	x	x	x	x	x	x	x					x		T
A50	Implementation cost	x	x	x	x	x	x	x	x					x		T
A51	Human resources cost	x	x			x	x	x	x					x		S
A52	Training cost													x	x	T
A53	Data quality cost	x	x			x				x				x		T
A54	Operations cost	x	x	x	x	x	x	x	x					x		T
A55	Systems cost	x	x	x	x	x	x	x	x					x		T
A56	Data collection cost	x	x											x		T
A57	Data storage cost						x							x		T
A58	Data processing cost					x								x		T
A59	Data retrieval cost	x		x			x		x					x		T
A60	Data monitoring cost	x	x	x	x		x							x		T
A61	Workload	x	x			x		x	x						x	S
A62	Lack of staff	x	x			x		x	x						x	S

Continued on next page

Continued from previous page

ID	Challenge	DC	DS	DT	DA	DC	DS	DA	DU	MDM	DLCM	DSPM	T/I	F/C	HC	S/T
A63	Staff skill level	x	x			x		x	x						x	S
A64	Lack of training	x	x			x		x	x						x	S
A65	Lack of participation	x	x			x		x	x						x	S
A66	Understand rationale of IS	x	x												x	S
A67	Group type context	x	x			x		x	x						x	S
A68	Health literacy	x	x			x		x	x						x	S
A69	Productivity	x	x			x		x	x						x	S
A70	Perception	x	x			x		x	x						x	S
A71	Self-efficacy	x	x			x		x	x						x	S
A72	Social influence	x	x			x		x	x						x	S
A73	Language barrier	x	x			x		x	x						x	S
A74	Basic education	x	x			x		x	x						x	S
A75	Digital illiteracy	x	x			x		x	x				x		x	S
A76	Lack of professionals	x	x			x		x	x			x			x	S
A77	Patient illiteracy	x	x												x	S
A78	Age issues	x	x			x		x	x						x	S
A79	Theft												x		x	S
A80	Ease of implementation												x			T
A81	Implementation completion												x	x		T
A82	Scalability									x			x	x		T
A83	Organisational structure and support									x			x	x	x	S;T
A84	Technical infrastructure available												x	x		T
A85	Maintenance												x	x		T
A86	Training													x	x	S
A87	Skilled personnel													x	x	S
A88	Adoption/ Resistance to change														x	S
A89	Stakeholder inclusion														x	S
A90	Behavioral/cultural aspect														x	S
A91	Local ownership									x					x	S

Appendix C

Knowledge transfer from existing maturity models

Knowledge from the existing maturity models was transferred with regard to:

1. the development of maturity models (Table C.1);
2. healthcare data collection (Table C.2);
3. data storage (Table C.3);
4. data sharing (Table C.4);
5. data analysis (Table C.5);
6. data usage (Table C.6);
7. data privacy and security (Table C.7);
8. data governance (Table C.8);
9. data technology and infrastructure (Table C.9);
10. human contributions (Table C.10); and
11. technological investments (Table C.11).

Table C.1: Knowledge transfer of the development of maturity models

Development of maturity models
Look at whole care pathway rather than placing a service at the centre of the evaluation (Flott *et al.*, 2016)
Consider the external challenges that affect maturity (Flott *et al.*, 2016)
Continued on next page

Continued from previous page
Development of maturity models
The maturity model should be multi-dimensional (Flott *et al.*, 2016) (Liu *et al.*, 2011) (Mettler and Blondiau, 2012) (Pearce *et al.*, 2013)
The stakeholders should be included in the model design process (Flott *et al.*, 2016)
Capability dimensions that can be considered in the model development are managing information, using business intelligence, using information technology, aligning business and information and managing change (Directorate Informatics)
Capability dimensions that can be considered in the new model are EMR, strategy, information security, data analysis, systems and IT infrastructure and people (Vidal Carvalho *et al.*, 2019)
Adopt a process focus rather than a function focus (Blondiau *et al.*, 2013)
Maturity should be assessed across different layers that may include a strategic (external cooperation of the organisation), organisational (internal or process) and information (IT) layer (Mettler and Blondiau, 2012) (Blondiau *et al.*, 2013)
Develop the maturity scale to be developer neutral and that has a broad applicability (Venesco, 2015)
Use design science research when developing the maturity model (Fitterer and Rohner, 2010)
Develop a comprehensive model with sufficient detail, with characteristics relating to maturity stages and tools for determining maturity (Carvalho *et al.*, 2019*a*)
Computerisation of data and management has three stages: collection, data sharing and data analysis (Sanders *et al.*, 2013)
Use a multi-faceted approach (combination) to measure maturity which may include process-oriented, people-oriented and technical aspects (Carvalho *et al.*, 2019*b*)
Develop a national framework to provide patient centred services to meet local needs (Johnston, 2017)
Develop a maturity model with processes that reaches up to a national level (Sharma, 2008)
Use design decisions when developing the maturity model (Mettler and Blondiau, 2012) (Carvalho *et al.*, 2019*a*)
Understand and use the 5 levels of CMMI which include initial, managed, defined, quantitatively managed, optimising (Fitterer and Rohner, 2010)
Iterative development process of maturity models (Carvalho *et al.*, 2019*b*)

Table C.2: Knowledge transfer of healthcare data collection

Data collection
Computerisation enables the automation of billing and accounts receivable (Liu *et al.*, 2011)
Continued on next page

Continued from previous page

Data collection
Cloud computing can enable the remote monitoring of chronic disease of patient at home (Grindle *et al.*, 2013)
EMR is a very important factor of healthcare data management (Vidal Carvalho *et al.*, 2019)
Data collection at facilities include clinical applications, ancillary services application, patient management applications and administrative applications (Jaana *et al.*, 2009)
Data collection is the first stage of computerisation of data and data management (Sanders *et al.*, 2013)
EMRs can facilitate the collection of population health data for analytics, planning, and delivery (Jones *et al.*, 2017)

Table C.3: Knowledge transfer of healthcare data storage

Data sources/Data storage
Cloud storage enables access of medical images anywhere and patients' images are available at different institutions (Grindle *et al.*, 2013)
EHR in the cloud should be a priority investment (Grindle *et al.*, 2013)
EMR and other structured data repositories are important (Johnston, 2017)
When implementing cloud computing it is important to keep some elements in in-house data centres that establish a hybrid IT storage structure (Grindle *et al.*, 2013)
EMR is important for mature data management (Rimmer *et al.*, 2014)
Listing different clinical applications, ancillary services application, patient management applications, administrative applications, clinical technologies and administrative technologies. Very extensive list
First start with widespread EMR adoption (Sanders *et al.*, 2013)
Enterprise Data Warehousing (big data) expands to include bedside devices, home monitoring data, external pharmacy data and detailed activity-based costing.
Their most important factors are EMR, strategy, information security, data analysis, systems and IT infrastructure and people. People are the greatest sub-area
EMRs are the main source of patient data - Mature to exhaustive EMRs for patients (Carvalho *et al.*, 2019*a*)
Cloud can enable HIE through connecting different EMR systems (Grindle *et al.*, 2013)

Table C.4: Knowledge transfer of healthcare data sharing

Data sharing
Interoperability is important for maturity (Flott *et al.*, 2016; Fitterer and Rohner, 2010; Naudet and Chen, 2012; Venesco, 2015)
Patient matching and data exchange is very important for effective healthcare delivery (Venesco, 2015)
The Internet of Everything can be used to connect data records and support digitalisation (Frost & Sullivan, 2015)
It is important that data is shared among members of the workflow team (Sanders *et al.*, 2013)
Data should be shared in a timely fashion with minimum latency (Sanders *et al.*, 2013)
The networkability of organisations should incorporate internal and external capabilities to collaborate with each other at the level of both business processes and underlying ICT infrastructure (Fitterer and Rohner, 2010)
Interoperability is the ability of two or more systems or components to exchange information and to use that information that has been exchanged (Naudet and Chen, 2012)
For a comprehensive model, all service providers should be included in the healthcare process (Sharma, 2008)
The maturity model should continuously improve entity integration, department integration and infrastructure development (Sharma, 2008)
Informatics provide quality information and analytics to inform business and clinical decision-making, streamline and automate processes and services, and provide safe and open access to information (Directorate Informatics)
Data sharing is the second stage of healthcare computerisation (Sanders *et al.*, 2013)

Table C.5: Knowledge transfer of healthcare data analysis

Data analysis
Business intelligence and data analytics should be included under data analysis (Carvalho *et al.*, 2019*a*)
DA is important for decision-making (Carvalho *et al.*, 2019*b*)
See data analytics maturity identifiers in HIS (Carvalho *et al.*, 2019*b*)
Business intelligence are the strategies and technologies used for data analysis of business information (Carvalho *et al.*, 2019*b*)
Data should be analysed to see patterns in analysed data and aggregated big data analysis (Sanders *et al.*, 2013)
Data analysis should be characterised by the adoption of enterprise data warehouses for big data analytics (Sanders *et al.*, 2013)
The maturity model should describe the progressive increase in complexity of analytic algorithms and data binding (Sanders *et al.*, 2013)
Continued on next page

Continued from previous page

Data analysis
Progress to holistic analysis solutions rather than fragmented point solutions that have very limited analytics capabilities (Sanders *et al.*, 2013)
Population based analytics can be used to improve individual patient care (Sanders *et al.*, 2013)
The healthcare analytical motive is towards wellness management, physical and behavioural-functional health and mass customisation of precise, patient tailored care (Sanders *et al.*, 2013)
Data analysis should include prescriptive analytics and interventional decision support (Sanders *et al.*, 2013)
Data analysis should mature to a state where further medical innovation for data analytics are enabled (Sanders *et al.*, 2013)
Data analysis is a very important factor of effective healthcare delivery (Vidal Carvalho *et al.*, 2019)
Senior decision-makers should be equipped in analytics (Johnston, 2017)
Analytics inform business and clinical decision making, streamline and automate processes and services, and provide safe and open access to information (Directorate Informatics)
Data analysis is the third stage of computerisation (Sanders *et al.*, 2013)
The Enterprise Data Warehouse must be agile for reporting due to external reporting requirements (Sanders *et al.*, 2013)

Table C.6: Knowledge transfer healthcare data usage

Data usage
Cloud computing can enabled virtual visits to patients (Grindle *et al.*, 2013)
Big data in the cloud can be harnessed to improve preventative well-being (Grindle *et al.*, 2013)
The analysed data of DA can be used for decision-making (Carvalho *et al.*, 2019*b*)
Data should be used to produce efficient, accurate and consistent reports that are available across the organisation (Sanders *et al.*, 2013)
Better integrated care
Use business intelligence and data analytics for decision-making (Carvalho *et al.*, 2019*a*)
Cloud computing can enable remote monitoring of chronic disease of patient at home (Grindle *et al.*, 2013)
Informatics and analytics can be used to inform business and clinical decision-making, streamline and automate processes and services, and provide safe and open access to information. (Directorate Informatics)

Table C.7: Knowledge transfer of healthcare data security and privacy management

Data security and privacy management
Security is important for maturity (Impact Advisors, 2015)
Security is enabled through technology, infrastructure and processes (Huang *et al.*, 2008)
Legislation, policies, standards and regulations are important for security and privacy (Huang *et al.*, 2008)
The main goals of data security are confidentiality, integrity and availability (Carvalho *et al.*, 2019*a*)
Data sharing through the cloud should be authorised (Grindle *et al.*, 2013)
Data security is a very important factor of effective healthcare delivery (Vidal Carvalho *et al.*, 2019)
Data security and privacy is a necessary component to enable effective digitalisation (Johnston, 2017)
Security standards are important for healthcare delivery (Liu *et al.*, 2011)

Table C.8: Knowledge transfer of healthcare master data management

Master data management
A digital architecture should be developed for digital data management (Johnston, 2017)
The data strategy should align to patient outcomes (Johnston, 2017)
The aspects that are important for data matching are data quality, processes and relevant regulations (Venesco, 2015)
Data elements should reduce duplication with supporting processes and regulations for data matching (Venesco, 2015)
Data governance function start with reducing organisational and cultural barriers to data access, increasing data quality in the source systems and master data identification and management. (Sanders *et al.*, 2013)
Identify and standardise vocabularies and reference data across disparate source systems (Sanders *et al.*, 2013)
Data strategy is a very important factor for the effective delivery of healthcare (Vidal Carvalho *et al.*, 2019)
Computerised medical records should be standardised for national benchmarking (Pearce *et al.*, 2013)
Data quality challenges and standardisation needs to be addressed to enable mature computerised medical records (Pearce *et al.*, 2013)
The governance of records and data repositories is important (Pearce *et al.*, 2013)
Data governance support multidisciplinary care management teams that focus on improving the health of patient populations (Sanders *et al.*, 2013)
SSoT should be maintained (Sanders *et al.*, 2013)

Table C.9: Knowledge transfer of healthcare data management technology and infrastructure

Technology and infrastructure
ICT infrastructure enables networkability within and between hospitals (Fitterer and Rohner, 2010)
E-health readiness includes technological components (Liu *et al.*, 2011)
IT infrastructure capability is important for transaction based applications (Liu *et al.*, 2011)
Systems and IT infrastructure is very important for the effective delivery of healthcare (Vidal Carvalho *et al.*, 2019)
Clinical technologies and administrative technologies should be implemented for healthcare data management (Jaana *et al.*, 2009)
Technology is used to contribute to the security of data(Huang *et al.*, 2008)
Healthcare data management infrastructure development can be improved through the use of a maturity model (Sharma, 2008)

Table C.10: Knowledge transfer of the human contributions in healthcare data management

Human contributions
Senior decision-makers should be equipped in analytics (Johnston, 2017)
When introducing change in a system, user resistance is important to take into account (Liu *et al.*, 2011)
Change management is very important in order to reach maturity (Jones *et al.*, 2017)
Maturity of a system is not only dependent on the installation of hardware and software, but dependent on effective change management too (Jones *et al.*, 2017)
The data literacy among employees is important for better data analysis (Sanders *et al.*, 2013)
Data governance function start with reducing organisational and cultural barriers to data access, increasing data quality in the source systems and master data identification and management. (Sanders *et al.*, 2013)
A sustainable data-driven culture should be achieved (Sanders *et al.*, 2013)
People are a very important factor to enable effective healthcare data management (Vidal Carvalho *et al.*, 2019)
E-health readiness includes managerial components (Liu *et al.*, 2011)

Table C.11: Knowledge transfer of healthcare data management finances and cost

Finances and cost
IT investments of EHR in the cloud should be a priority (Grindle *et al.*, 2013)

Appendix D

Model development iterations

This appendix summarises all the changes that were made to the model at the different iterations. The contributions of different SMEs at different iterations are also indicated.

Table D.1: Model iterations

#	Main changes wrought	Maturity levels	Maturation paths	System levels	Domain components	# of capability areas	Inputs for changes
1	Initial structure of model with capability areas that are merged and not distinct	Unspecified (different number for each domain component	Maturation path dependent on domain component	Unspecified	Data collection, Data sharing, Data analysis	11	Literature
2	Expansion on different capability areas, more distinct capability areas versus merged ones, include system levels in the model, break up merged domain components for more distinct domain components, include data storage and data usage domain components	Unspecified (different number for each domain component	Maturation path dependent on domain component	Facility, Organisation	Data collection, Data storage, Data sharing, Data analysis, Data usage	60	Literature
3	Defined the maturation path across the different domain components, all domain components follow the same maturation path and has the same maturity levels, exclude some redundant capability areas from existing ones	Level 1: Ad hoc paper based, Level 2: Initial paper based, Level 3: repeatable electronic, Level 4: Defined electronic, Level 5: Managed electronic, Level 6: Optimising electronic	One maturation path applied across the different domain components	Facility, Organisation	Data collection, Data storage, Data sharing, Data analysis, Data usage	57	Literature
4	Include data governance domain component with initial data governance capability areas, incorporated different maturation paths, expand more capability areas and descriptions of existing capability areas	Level 1: Ad hoc paper based, Level 2: Initial paper based, Level 3: repeatable electronic, Level 4: Defined electronic, Level 5: Managed electronic, Level 6: Optimising electronic	Different maturation paths for primary activities, support structures and enabling practices for all domain components	Facility, Organisation	Data collection, Data storage, Data sharing, Data analysis, Data usage, Data governance	69	Literature

Continued on next page

Continued from previous page

#	Main changes wrought	Maturity levels	Maturation paths	System levels	Domain components	# of capability areas	Inputs for changes
5	Defined capabilities that are on both the facility and organisational level and not as mutually exclusive facility and organisational levels	Level 1: Ad hoc paper based, Level 2: Initial paper based, Level 3: repeatable electronic, Level 4: Defined electronic, Level 5: Managed electronic, Level 6: Optimising electronic	Different maturation paths for primary activities, support structures and enabling practices for all domain components	Facility, Organisation	Data collection, Data storage, Data sharing, Data analysis, Data usage, Data governance	73	Literature
6	Include data processing capability area under data collection, include business rules under data governance, include adherence to governmental acts capability area, changed definition of maturity concept to process maturity (how well a capability is carried out without referring to paper-based or electronic systems), changed the number of maturity levels, changed maturity level definitions, exclude redundant capability areas	Level 1: Initial, Level 2: Repeatable, Level 3: Defined Level 4: Managed, Level 5: Optimising	Different maturation paths for primary activities, support structures and enabling practices for all domain components	Facility, Organisation	Data collection, Data storage, Data sharing, Data analysis, Data usage, Data governance	71	Literature, SME 1: Maturity models SME 2: SQL, enterprise architecture

Continued on next page

Continued from previous page

#	Main changes wrought	Maturity levels	Maturation paths	System levels	Domain components	# of capability areas	Inputs for changes
7	Validate captured data, include improvement to automated processes in optimising stage, include addressing different scenarios for data life cycle management, include the right of sharing data in adherence to government acts, infrastructure should not be limited to physical infrastructure, but can mature to alternative structures, include the security of devices, expand data rules with labelling data for different categories, include the data architecture in the data strategy, expand description of data queries, include Master Data Management capability area, combine data sending, receiving and interoperability, change wording to describe the capability of structured and unstructured data storage, include ethical considerations for big data capturing, include data processing for sharing, change data analysis descriptions on organisational level, change descriptions of data sharing on both the facility and organisational level	Level 1: Initial, Level 2: Repeatable, Level 3: Defined Level 4: Managed, Level 5: Optimising	Different maturation paths for primary activities, support structures and enabling practices for all domain components	Facility, Organisation	Data collection, Data storage, Data sharing, Data analysis, Data usage, Data governance	77	Literature, SME 3: Data engineering SME 4: Data engineering

Continued on next page

Continued from previous page

#	Main changes wrought	Maturity levels	Maturation paths	System levels	Domain components	# of capability areas	Inputs for changes
8	Include data ingestion, include the capturing and storage of financial data for patient care, include the capturing and storage of medicinal inventory and stock orders data, include workforce management and human capital data capturing and storage, refinement of all maturity level descriptions to make it more concise and distinct, defined domain concepts and unspecific terminology, moved data sharing technology and infrastructures capability areas, added facility level usage capability areas utilisation management of infrastructure and medical stock level management, development of the transfer media with an introduction, overview of system levels, instructions to do assessments, assessment sheets per domain component and results sheets, included ethical considerations more explicitly under data governance domain component	Level 1: Initial, Level 2: Repeatable, Level 3: Defined Level 4: Managed, Level 5: Optimising	Different maturation paths for primary activities, support structures and enabling practices for all domain components	Facility, Organisation	Data collection, Data storage, Data sharing, Data analysis, Data usage, Data governance	87	Literature, Verification SME 1: Data scientist employed in healthcare sector, Verification SME 2: Process improvement engineer employed
9	Include data collection of health indicators, include the specification of health indicators, include metadata management, eliminate data collection for management and decision-making, refine data storage for management and decision-making to data aggregation and storage for management and decision-making, include contractual component and consent considerations to the alignment to storage and sharing policies, standards and regulations, include the assistance of third-parties for maturity assessments, the inclusion of result indicators to assist users interpret results and identify the components to improve on, improvement on the structuring and layout of the introductory and overview sheets	Level 1: Initial, Level 2: Repeatable, Level 3: Defined Level 4: Managed, Level 5: Optimising	Different maturation paths for primary activities, support structures and enabling practices for all domain components	Facility, Organisation	Data collection, Data storage, Data sharing, Data analysis, Data usage, Data governance	89	Validation SME 1, Validation SME 2, Validation SME 3, Validation SME 4, Validation SME 5, Validation SME 6, Validation SME 7

Appendix E

The HCDMMM capability areas maturity level descriptions

Table E.1 gives all the maturity level descriptions of all the capability areas. In Table E.1 the following keys are important:

- Facility (F)
- Organisation (O)
- Capability Area (CA)
- Capability category (CC)
- Enabling Practices (EP)
- Primary Activity (PA)
- Support Structure (SS)
- System level (SL)

Table E.1: Capability area maturity level descriptions

Collection							
SL	**CA**	**CC**	**1. Initial**	**2. Repeatable**	**3. Defined**	**4. Managed**	**5. Optimising**
F	Data capturing for patient care	PA	Patient data is captured with initial methods and procedures to be followed. Data capturing is time-consuming and is seldom captured accurately and effectively	Data capturing can be done routinely with repeatable accuracy. Capturing is controlled through simple and basic methods and procedures that address the full intent of the data capturing practices	The methods and procedures of data capturing are well-defined for consistent and effective data capturing. The different methods and procedures for different data collection are well-documented. Data capturing methods and procedures function as a coherent whole	Methods and procedures of patient data capturing are supervised and managed. Accuracy and consistency of different data capturing processes are monitored to detect the area of focus for corrective action in order to improve data capturing	The range of ability of data capturing is improved proactively through exploiting data capturing best practices innovatively. Data capturing processes are continuously improved to be carried out automatically
F	Data collection of health indica-tors	PA	The facility is not aware of the health indicators that it must collect	The facility is aware of the health indicators specified by the organisation that it must collect. Collection is controlled through simple and basic methods and procedures to collect health indicators	The facility effectively collects the core set of health indicators that the organisation specified which includes health status, risk factors, service coverage and health systems. The facility understands the importance of collecting health indicators and a common understanding of responsibilities are established	The facility monitors changes in the list of core health indicators specified by the organisation. The facility checks completeness of health indicator reports sent to the organisational level and takes corrective action when needed	The collection of core health indicators are continuously optimised to minimise the workload of healthcare workers that collect the data. The report checks are continuously optimised to identify problem areas and to address them proactively
F	Financial data capturing of patient care	PA	Financial data for care delivery is captured with initial methods and procedures to be followed. Data capturing is time-consuming and is seldom captured accurately and effectively	Financial data capturing for care delivery is controlled through simple and basic methods and procedures. Financial data can be accurately captured, but is resource-intensive	The methods and procedures of data financial data capturing are well-defined for consistent and effective data capturing. Data capturing of financial data functions as a coherent whole with care delivery	Methods and procedures of financial data capturing for care delivery are supervised and managed. Accuracy and consistency of different data capturing processes are monitored to detect the area of focus for corrective action	Financial data capturing of care delivery is improved proactively through exploiting data capturing best practices of financial data innovatively. Data capturing processes are continuously improved to be carried out automatically

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Medicinal inventory and stock orders data capturing	PA	No defined methods or procedures exist to capture inventory and stock orders data	Basic techniques, methods and procedures are defined and used to capture medicinal inventory and stock order data. Data capturing is resource-intensive and not always accurate	Effective and efficient medicinal inventory and stock order data capturing methods and procedures are established and well-documented	Methods and procedures of medicinal inventory and stock order data capturing are supervised and managed. Accuracy and consistency of different data capturing processes are monitored to detect the area of focus for corrective action in order to improve data capturing	The range of ability of data capturing is improved proactively through exploiting data capturing best practices innovatively. Data capturing processes are continuously improved to be carried out automatically
F and O	Validate captured data	PA	Captured data is not validated	The captured data is validated repeatedly, resulting in the elimination of errors in the captured data. Data validation is controlled through basic methods and procedures. Validation of captured data is time-consuming, is dependent on dedicated individuals and seldom results in validated data that eliminates errors like missing data, data duplication and other inconsistencies effectively	All the different kinds of data validation methods and procedures are well-defined and implemented for effective data validation. The inputs, standards and procedures for validating data are well-defined and implemented. Data validation results in minimal errors in the data	Data validation methods and processes that do not result in effectively validated data are detected and refined to further improve data validation	The range of ability, methods and procedures of the validation of captured data are continuously improved by innovatively exploiting data validation best practices. Data validation continuously improves the automatic data validation of all captured data

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Processing of collected data	PA	Captured data is not processed to ensure it is useable	Basic captured data processing methods and procedures are repeatable. Raw data is transformed to be useable and easily interpreted, but the processing is resource intensive and time-consuming	The data processing methods and procedures are well-documented. The processing methods and procedures are effective at making data easily understandable and useable. Processing is standardised and consistent	Data processing methods and processes that do not result in easily understandable and useable data are detected and refined through corrective measures that are in place	The processing of data is continuously improved by exploiting best data processing practices innovatively. Collected data processing continuously improves towards the automatic data processing of all captured data
F and O	Data collection technology	SS	Technologies for data collection exist. The available technologies gives very limited support to accomplish effective data collection	Data collection technologies are adequate, robust, reliable and available. Adequate hardware and software are available for repeatable data collection	Data collection technology is compatible and interoperating with other devices and applications. Technologies for structured and unstructured electronic data collection are applied	Availability, reliability and maintainability of data collection technology are monitored and corrective maintenance is executed effectively and timely when necessary	Preventative maintenance and upgrades of data collection technology are executed proactively, timely and effectively. New innovative technologies are continuously implemented
F and O	Data collection infrastructure	SS	Infrastructure for data collection exists, but gives very limited support to the effective execution of data collection	Appropriate infrastructure for data collection is applied that gives adequate support for repeatable data collection	Specific infrastructure for data collection is applied that supports data collection effectively. Infrastructure for structured and unstructured data collection are applied	The availability, reliability and maintainability of data collection infrastructure are monitored. When failure occurs infrastructure is repaired or replaced	Data collection infrastructure is continuously maintained and upgraded whenever needed. Infrastructure is continuously improved towards new innovative or alternative structures
F and O	Data entry forms / structure	EP	Initial data entry forms do not enable the capturing of all relevant data. Entry forms are not easy to use and important data is left out	Data entry forms enable the adequate and consistent capturing of relevant data.	Well-established data entry forms enable the effective capturing of all relevant data. The data entry forms are easy to use, intuitive and not time-consuming. The data entry forms are standardised across the different points of care	The data entry forms are controlled and checks are in place to ensure appropriate data entry onto the forms across the different care points	The data entry forms for data capturing are continuously optimised to enable the most effective data capturing. The ease to enter data and time it takes are continuously improved. Measures to ensure data is correctly entered are continuously improved

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Comprehensiveness of data captured	EP	Data is being collected without the adequate knowledge of what all the necessary data elements are for baseline care delivery and decision-making	Data is being collected with the adequate knowledge of what all the necessary data elements are for baseline care delivery and decision-making	Comprehensive data is being collected, according to the extensive knowledge of what all the necessary data elements are, for effective care delivery and decision-making	It is monitored whether the comprehensive sets of data are effectively applied according to the purpose that they were collected for. Corrective action is taken when data is not effectively applied or collection of data is discontinued if redundant	The comprehensiveness of data is continuously improved when new applications for data not being collected yet surface for optimised care delivery and decision-making
F and O	Types of data captured	EP	Paper-based and electronic data are captured. Data is not captured in a controlled way. Vast amounts of data is captured on paper	Paper-based and electronic data is captured. Electronic data collection consists of structured data only	Paper-based and electronic data structures are defined. Electronic data types that are collected include structured and unstructured data	Paper-based, structured and unstructured electronic data capturing are managed effectively. If paper-based capturing is done, it is because it is more effective than electronic capturing for its purpose and not because of legacy systems	The various types of paper-based and electronic data that are captured are continuously improved for the efficient and effective application of the collected data according to its purpose
F and O	Quality of collected data	EP	Collected data is of varying and unacceptable quality	Data of similar quality is consistently collected, but no defined standard of quality data exist	The designated role and purpose, consistency and timeliness of collected data are defined and standardised. An entity-wide understanding of the quality of data metrics exist	Data is collected effectively, consistently and timely according to its designated role and purpose	Causes of unacceptable quality data is continuously and effectively identified and addressed
F and O	Data collection privacy and security	EP	Collected data is not secured and patient privacy is not guaranteed	Basic authorised access control to collected data exist to ensure the availability, integrity and confidentiality of healthcare data. Access options between view data, insert data, update data and delete data is differentiated for privacy and security	Security software is in place to protect collected data against malware and procedures are in place to prevent malicious use of data by authorised users. Operational and management security for data collection are well-defined.	Confidentiality and integrity of collected data are monitored to ensure data stays available to only the intended users and to ensure data is not corrupted or stolen. Measures are in place to identify malicious use of data through authorised users	Continuously upgrade and update data collection privacy and security software. Security and privacy procedures and checks are continuously improved to ensure availability, integrity and confidentiality of the collected data

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Security of physical data collection devices	EP	Data collection devices are not secured. Collection devices can easily be stolen that prevents effective collection of data	Basic security policies and procedures are in place for the security of physical data collection devices. Security policies and procedures does not completely ensure devices are secure and not stolen	Security policies and procedures for physical data collection devices are well-defined and established that ensure a stable environment for data collection through the devices	The security of physical data collection devices are monitored. The correct usage of the devices are monitored to ensure it is not used for the wrong intent. Users that use the devices for the wrong reasons can be identified	The security of physical data collection devices are continuously improved. Device security policies and methods are continuously updated to prevent new threats and misuses of the devices
F and O	Align-ment to data collection stan-dards, policies and regu-lations	EP	Identified need for the alignment to data collection standards, policies and regulations. No standards, policies or regulations are implemented or consistently aligned to	No alignment to data collection standards, policies and regulations. If they are aligned to, it is incidental and unpredictable	The alignment to specific data collection standards, policies and regulations ensures the effective functioning of data collection. The standards, policies and regulations ensure that an organisation-wide understanding of activities, roles and responsibilities are in place	There are alignment to standards, policies and regulations control measures in place like auditing and alignment monitoring that allows for quantitative feedback	Complete alignment to data collection standards, policies and regulations are carried out that enable strategy realisation. Any issues of non-alignment are identified and remedial action is taken to ensure alignment in a timely manner
O	Work-force and human capital data collection	PA	No clear procedures and processes exist to collect workforce and human capital data. Data is seldom collected. Collected data is disorganised	Workforce and human capital data is collected following basic defined processes and procedures. Data is repeatedly collected and collection is organised. Different facilities' data are in silo's. A basic set of workforce and human capital data is collected	Workforce and human capital data is collected following standardised procedures and processes across the whole organisation. Collected data is consolidated. The data that needs to be collected is standardised across all organisational entities. Workforce and human capital data is collected for specific purposes	The execution of the collection of workforce and human capital data is supervised and monitored. Areas of ineffective data collection can be detected and corrective action is taken. The skills, knowledge and experience of the workforce are collected	What workforce and human capital data to collect and the collection procedures and processes are continuously optimised in order to determine indicators like the supply, demand and distribution of the workforce
Storage							

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Patient data storage into repositories	PA	Initial methods and procedures are followed to store patient data. Storage is resource-intensive and is carried out manually. The storage of data is not organised or controlled	Data from different types of care within the F is stored in different data repositories. Data storage repositories are organised and controlled and data can be stored following repeatable methods and procedures. Structured and unstructured data are stored in the appropriate repositories	The methods and procedures of data storage are well-documented and effective. The different data repositories are integrated into a coherent whole. Structured data is effectively stored in the appropriate repositories for structured data and unstructured data is effectively stored in appropriate repositories for unstructured data	The efficiency and effectiveness of the data storage are monitored through meaningful and well-established performance metrics. Inefficient data storage methods and procedures are identified and corrective action is taken timely and effectively to ensure data storage efficiency and effectiveness	Structured data repositories are optimised for control of structured data and unstructured data repositories are optimised for access and flexibility. The range of the ability of data repositories are continuously improved and storage repositories are continuously improved towards automated storage
F	Financial data storage for care of patients	PA	Initial methods and procedures are followed to store financial data. Storage is resource-intensive and is carried out manually. The storage of financial data is not organised or controlled	Financial data from different types of care within the F is stored in different data repositories. Data storage repositories are organised and controlled and data can be stored following repeatable methods and procedures	The methods and procedures of data storage of financial data are well-documented and effective. The different financial data repositories are integrated into a coherent whole and managed centrally	The efficiency and effectiveness of the data storage of financial data are monitored through meaningful and well-established performance metrics. Inefficient data storage methods and procedures are identified and corrective action is taken timely and effectively	Financial data storage repositories are continuously improved towards automated storage. Monitoring methods are improved to ensure financial data is correct
F	Inventory and stock orders data storage	PA	Initial methods and procedures are followed to store inventory and stock orders data. Storage is resource-intensive and is carried out manually. The storage of data is not organised or controlled	Inventory and stock orders data for different types of care within the facility is stored in different data repositories. Data storage repositories are organised and controlled and data can be stored following repeatable methods and procedures	The methods and procedures of data storage of inventory and stock orders data are well-documented and effective. The different inventory and stock order data repositories are integrated into a coherent whole and are managed centrally	The efficiency and effectiveness of the data storage of inventory and stock orders data are monitored through meaningful and well-established performance metrics. Inefficient data storage methods and procedures are identified and corrective action is taken timely and effectively	Inventory and stock orders data storage repositories are continuously improved towards automated storage. Inventory and stock orders data is automatically updated. Monitoring methods are improved to ensure inventory and stock orders data is correct

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Data ingestion	PA	Data ingestion is very resource intensive, not stable or controlled. Effective data ingestion from collection to storage is not the norm and is sporadic. Ingestion errors often occur	Data ingestion from collection to storage is done repeatedly following a disciplined way which is stable and controlled, but resource intensive	The data ingestion processes, methods and procedures are well-documented and effective. Data ingestion is not resource intensive. Well-established processes and well-defined methods and procedures ensure effective ingestion of data from collection to storage	Execution of data ingestion methods and processes are auditable to maintain stable and effective data ingestion. Because the data ingestion process is auditable, it is possible to take corrective action when stored data does not correlate with collected data	Collected data is automatically and immediately imported into its appropriate storage repository for immediate use. As it is collected, it is immediately available in its appropriate storage repository. Data sources are prioritised and data items are routed to the correct destinations. The method and range of ability of data ingestion is continuously improved
F and O	Data storage backups	PA	No formal data backup solution is officiated	Backup solutions backup data repeatedly at pre-determined frequencies by copying it to a different location, but is resource-intensive	The critical data to be backed up is documented by a backup administrator. The backup system is effective and backs up data as a coherent whole. The backups are consistent and of good quality	The backup system is monitored through well-established backup performance metrics. The back up scope, schedule, recovery point and recovery time objectives are managed effectively	Techniques for efficient and cost-effective backups are continuously optimised. Ineffective back up data storage, computation and network resources are continuously identified and improved
F and O	Data quality processes for stored data	PA	Stored data is not processed to ensure good quality. Quality is not consistent	Basic and repeatable procedures and processes to check and improve the quality of stored data are applied. Data quality processes ensure consistent data	The data quality standards for stored data are defined and stored data is processed effectively to ensure high quality of stored data	The quality processes and procedures for stored data that do not result in good quality data are detected and refined thorugh corrective measures that are in place	The data quality processes for stored data are optimised to continuously monitor, check and improve the quality of stored data automatically

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Data storage technology	SS	The data storage capacity is low and limited. The different types and formats of data that can be stored is limited. Storage technology cannot guarantee data availability, security and reliability. Storage utility is inefficient and storage is not easily scalable. Storage technology can handle very limited read and write operations and it takes a lot of time to complete an operation	Data storage technologies are adequate, reliable and available. Storage capacity is adequate for the storage activities. Data availability and security are mostly guaranteed. Storage utility is adequate. Storage technology can handle adequate read and write operations and it takes adequate time to complete an operation	The data storage capacity is high, many different data types and formats can be stored and storage is easily scalable. Storage utility is efficient. Storage technology effectively guarantees data availability, security and reliability. The storage technology read and write ability effectively and efficiently support data storage activities. Data storage technology is compatible and interoperating with other devices and applications	The performance metrics of the data storage technology are monitored. Performance metrics like storage capacity, Input/Output Operations Per Second and latency are monitored. Availability, reliability and maintainability of data storage technology are monitored and corrective maintenance is executed when necessary	Upgrades of data storage technology are executed timely and effectively. Technologies to store structured and unstructured data are continuously upgraded and updated. New innovative technologies are continuously implemented
F and O	Data storage infrastructure	SS	Existing infrastructure for data storage gives very limited support to the effective execution of data storage	Appropriate data storage infrastructure is applied that gives adequate support for repeatable data storage	Specific infrastructure for data storage is applied that supports data storage effectively. Infrastructure for structured and unstructured data storage are applied	The availability, reliability and maintainability of data storage infrastructure are monitored. When failure occurs infrastructure is repaired or replaced	Data storage infrastructure is continuously maintained and upgraded whenever needed. Infrastructure is continuously improved to new innovative or alternative structures

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Retention and disposal plan of stored data	EP	Initial plan for data retention and when to dispose obsolete or dormant data exist. Data anomalies occur with the deletion of stored data	Retention and disposal plan of obsolete or dormant data includes repeatable methods of data disposal stored in various forms. Paper-based data is shredded or incinerated. DVDs, CDs, hard disks and other electronic storage are overwritten or physically destroyed. The deletion of data causes very little data anomalies and data integrity is mostly maintained	Defined standards for disposal of obsolete or dormant data are applied. Data disposal is effective and does not cause data anomalies or data integrity issues. A defined plan of what, when and how to dispose data is followed. Retention and disposal of data from different scenarios and of various types are well-defined to retain or dispose of it according to specifications	The retention and disposal of data are well-managed. All overwritten or physically destroyed obsolete or dormant electronic data are referenced and registered. Data retention and disposal are managed according to specifications and corrective action is immediately taken when data retention and disposal varies from standards	Data retention and disposal are optimised in terms of minimal waste of storage space and cost of out dated data and to ensure active data is not disposed of too early
F and O	Storage privacy and security	EP	Stored data is not secured and patient privacy is not guaranteed	Data storage privacy and security measures are applied for consistent data security and privacy of stored data. Basic authorised access to devices and servers is implemented to prevent unauthorised access, modification and deletion of stored data	Effective privacy and security software and procedures are implemented to prevent access of unauthorised users, malicious behaviour of authorised users and to prevent malware from corrupting or stealing stored data	Confidentiality and integrity of stored data are monitored to ensure data stays available to only the intended users and to ensure data is not corrupted or stolen. Measures are in place to identify malicious use of data through authorised users	Storage system privacy and security software and control measures are continuously upgraded and updated to ensure the availability, integrity and confidentiality of stored data

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Alignment to data storage standards, policies and regulations	EP	Identified need for the alignment to data storage standards, policies and regulations. No standards, policies or regulations are implemented or consistently aligned to	No alignment to data storage standards, policies and regulations. If they are aligned to, it is incidental and unpredictable	The alignment to specific data storage standards, policies and regulations ensures the effective functioning of data storage. The standards, policies and regulations ensure that an organisation-wide understanding of activities, roles and responsibilities are in place. The necessary contractual and consent measures specified by relevant regulations and acts for data storage are followed	There are alignment to standards, policies and regulations control measures in place like auditing and alignment monitoring that allows for quantitative feedback	Complete alignment to data storage standards, policies and regulations are carried out that enable strategy realisation. Any issues of non-alignment are identified and remedial action is taken to ensure alignment in a timely manner
O	Data aggregation and storage for management and decision-making	PA	No clear methods or procedures are followed to store data aggregated from operational processes like clinical, financial or administrative processes for management and decision-making. Data storage is resource-intensive and is seldom stored effectively	Storage of data aggregated from operational processes like clinical, financial or administrative processes is controlled through initial and basic methods and procedures. Data for decision-making and management is stored routinely	All data aggregated from operational processes like clinical, financial or administrative processes for management and decision-making is stored effectively. The methods and procedures for the data storage for management and decision-making are well-defined and documented	The effectiveness of data storage for management and decision making of data aggregated from operational processes like clinical, financial or administrative processes and its methods and procedures are monitored through well-established data storage performance metrics	Best practices are exploited innovatively to store management and decision making data as effectively as possible and to continuously improve the range of ability of data-driven management and decision-making. Storage of data for management and decision-making are continuously improved towards automated aggregation and storage

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Storage of work-force and human capital data	PA	No clear procedures and processes exist to store workforce and human capital data. Collected data is stored in a disorganised way	Workforce and human capital data is stored following basic defined processes and procedures. Data is repeatedly stored and storage is organised. Different facilities' data are stored in silo's	Workforce and human capital data are stored following standardised procedures and processes across the whole organisation. Data is stored centrally. The data that needs to be stored is standardised across all organisational entities	The execution of the storage of workforce and human capital data is supervised and monitored. Areas of ineffective data storage can be detected and corrective action is taken	The storage procedures and processes of workforce and human capital data are continuously optimised in order to have the right data available for purposes like the supply, demand and distribution of the workforce
Sharing							
F	Data accessi-bility of different data reposito-ries within the facility	PA	Initial methods and processes takes long to access data from different repositories within the F and make data available at different care points in the F. The different data repositories are in silos and not linked	Data from different repositories within the F can be accessed repeatedly through basic methods and procedures. The data access methods and procedures can be carried out consistently	Methods and processes to access data from different repositories in the F are effective. Different queries of structured and unstructured data within the F are effective. Data can be queried and is easily accessible to defined authorised users	The structured and unstructured data that are readily available within the F are effectively managed through defined security and authorisation measures. Data is automatically available, but access must be approved for different authorised users before they can access it	The data accessibility from the different repositories in the F is continuously improved. Data sharing queries are optimised for flexible user views for the purpose of efficient data usage. Data accessibility is continuously improved to be immediately available to authorised users

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Data availability from and to the facility (facility interoperability)	PA	Data is sent and received from the facility to make data available across different locations. Data is sent and received infrequently and sending and receiving of structured and unstructured data are ineffective	Data is frequently and repeatedly sent and received from the facility through basic methods and procedures to make data available across different locations. Sent and received data are often in different formats and the systems are not interoperable. Basic procedures are established to make structured and unstructured data from and to different facilities exchangeable and usable	Structured and unstructured data are readily available to and from the facility. The effective interoperability of the facility between other facilities are established through standardised operating frameworks and protocols. The facility interacts with others through local and wide-area networks to function as a coherent whole with them	The structured and unstructured data that are readily available to and from the facility are effectively managed through defined security and authorisation measures. Data is automatically available, but access must be approved for different authorised users of facilities before they can access it	The methods and procedures to make data available to and from the facility are continuously improved to ensure data is securely available and accessible. The interoperability methods are continuously improved so that data is automatically exchanged and usable
F	Technology for data accessibility of different data repositories within the facility	SS	Technologies for data accessibility of different data repositories within the facility exist. The available technologies gives very limited support to accomplish effective data accessibility	Technologies for data access to different data repositories in the facility are available, robust and reliable. The proper hardware and software are available to establish access configurations for repeatable access between the different data repositories	Data access technologies are compatible and interoperable with other devices and applications in the facility. Technologies enables the effective queries between different repositories to support the data accessibility activities to function as a coherent whole	Reliability and efficiency of data access technologies are monitored and corrective measures are executed effectively and timely when technologies fail	Preventative measures and upgrades of data access technologies are executed timely and effectively to ensure that accessibility technologies continues to effectively support the accessibility of data in the facility. Technologies for accessing structured and unstructured data from their respective repositories of the facility are continuously upgraded and updated

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Infrastructure for data accessibility of different data repositories within the facility	SS	Existing infrastructure for data accessibility of different data repositories within the facility gives very limited support to the effective execution of data accessibility in the facility	Appropriate data accessibility infrastructure is applied that gives adequate support for repeatable data accessibility in the facility	Specific infrastructure for data accessibility in the facility is applied that supports structured and unstructured data accessibility effectively	The availability, reliability and maintainability data access infrastructure is monitored. When failure occurs, infrastructure is repaired or replaced	Data access infrastructure is continuously maintained and upgraded whenever needed to continue the support of data queries and accessibility in the facility. Infrastructure is continuously improved to new innovative or alternative structures
F	Technology for facility interoperability	SS	Technologies for data accessibility between the facility, other facilities and the organisation exist. The available technologies gives very limited support to accomplish effective data accessibility	Technologies for data access between the facility, other facilities and the organisation are available, robust and reliable. The proper hardware and software are available to establish access configurations for repeatable access	Data access technologies are compatible and interoperable with other devices and applications of other facilities and the organisation. Technologies enables the effective queries between the facility, other facilities and the organisation to support the data accessibility activities to function as a coherent whole	Reliability and efficiency of data access technologies are monitored and corrective measures are executed effectively and timely when technologies fail	Preventative measures and upgrades of data access technologies are executed timely and effectively to ensure that accessibility technologies continues to effectively support the accessibility of data between the facility, other facilities and the organisation. Technologies for accessing unstructured data and querying structured data from the respective facilities or the organisation are continuously upgraded and updated

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Infrastructure for facility interoperability	SS	Existing infrastructure for data accessibility between the facility, other facilities and the organisation gives very limited support to the effective execution of data accessibility	Appropriate data accessibility infrastructure is applied that gives adequate support for repeatable data accessibility between the facility, other facilities and the organisation	Specific infrastructure for data accessibility between the facility, other facilities and the organisation is applied that supports data accessibility effectively. Infrastructure for structured and unstructured data accessibility in the facility are applied	The availability, reliability and maintainability of electronic data access infrastructure is monitored. When failure occurs infrastructure is repaired or replaced	Data access infrastructure is continuously maintained and upgraded whenever needed to continue the support of data queries and accessibility from the respective facilities or the organisation. Infrastructure is continuously improved to new innovative or alternative structures
F and O	Data processing for data sharing	PA	Stored data is processed for sharing with no clear methods or procedures to be transformed into shareable formats. Processing is resource-intensive and ineffective. Processed data is not necessarily usable to the entities it is shared with	Stored data is processed with basic and repeatable methods and procedures to be shareable to different entities. Different processing methods and procedures exist for different sharing purposes	Different processing methods and procedures are well-established and documented for all the different sharing purposes so that all data is effectively shareable to different entities	Data processing for data sharing is monitored and controlled in accordance with the defined standards. Errors and anomalies in the processed data can be checked and corrective measures are in place	The effectiveness of processing stored data for sharing is continuously improved towards the automatic data processing of all data for sharing
F and O	Data transmission networks	SS	Initial data transmission networks are established to share data. Networks are slow and unreliable and gives limited support to the effective sharing of data	Appropriate data transmission networks exist to adequately support the sharing of data repeatedly and consistently	Specific data transmission networks are applied that effectively support sharing of data. Networks are capable of transmitting structured and unstructured data to their respective destinations	The capacity, configurations and channels of the data transmission networks are measured and monitored to identify causes for slow and ineffective data sharing through the data transmission networks. Corrective action is taken to ensure the networks support data sharing effectively	The capability, reliability and availability of data transmission networks are continuously optimised to support the most effective sharing of data. Continuous maintenance prevents failure of networks and ensure networks are never down so that data can be shared continuously. New innovative data transmission networks are continuously implemented

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Data sharing privacy and security	EP	Initial data sharing security and privacy checks and procedures applied cannot consistently maintain the confidentiality, availability and integrity of data being shared	Basic data sharing privacy and security measures across the data sharing networks of the F are applied for consistent security and privacy of data being shared. Authorised access to data and servers is implemented to prevent unauthorised sending and receiving of data	Effective data sharing privacy and security software and procedures are applied to prevent data leakage. The confidentiality of structured and unstructured data are ensured when transmitted. Networks are effectively secured for data sharing	Control measures are used to control network traffic into and out of the inter-connected networks. Potential damage of malware is effectively contained and attempts to break into servers and access data are effectively prevented. The confidentiality and integrity of data being transmitted are being monitored to ensure the secure transmission of data	Data sharing privacy and security software and control measures are continuously upgraded and updated to ensure the availability, integrity and confidentiality of data being transmitted
F and O	Security of data sharing devices	EP	Data sharing devices are not secured to enable the stable sharing of data. Sharing devices can easily be stolen that prevents effective sharing of data	Basic security policies and procedures are in place for the security of data sharing devices. Security policies and procedures does not completely ensure devices are secure and not stolen	Security policies and procedures for data sharing devices are well-defined and established that ensure a stable environment for data sharing through the devices	The security of data sharing devices are monitored. The correct usage of the devices are monitored to ensure it is not used for the wrong intent. Users that use the devices for the wrong reasons can be identified	The security of data sharing devices are continuously improved. Device security is continuously updated to prevent new threats and misuses of the devices

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Alignment to data sharing standards, policies and regulations	EP	Identified need for the alignment to data sharing standards, policies and regulations. No standards, policies or regulations are implemented or consistently aligned to	No alignment to data sharing standards, policies and regulations. If they are aligned to, it is incidental and unpredictable	The alignment to specific data sharing standards, policies and regulations ensures the effective functioning of data sharing. The standards, policies and regulations ensure that an organisation-wide understanding of activities, roles and responsibilities are in place. The necessary contractual and consent measures specified by relevant regulations and acts for data sharing are followed	There are alignment to standards, policies and regulations control measures in place like auditing and alignment monitoring that allows for quantitative feedback	Complete alignment to data sharing standards, policies and regulations are carried out that enable strategy realisation. Any issues of non-alignment are identified and remedial action is taken to ensure alignment in a timely manner
O	Data availability between the organisation and the facilities (organisation interoperability)	PA	Data is sent and received between the organisation and the facilities to make data available across different locations. Initial procedures for sending and receiving structured and unstructured data are ineffective and data is sent and received infrequently	Structured and unstructured data is frequently sent and received between the organisation and the facilities to make data available across different locations. Basic, repeatable sending and receiving methods and procedures are in place. Sent and received data are often in different formats and the systems are not interoperable. Basic procedures are established to make data between the organisation and different facilities exchangeable and usable	Structured and unstructured data are readily and automatically available between the organisation and the facilities. The effective interoperability of the organisation and the facilities are established through standardised operating frameworks and protocols for the effective exchange and use of data between locations. The organisation interacts with the facilities through local and wide-area networks to function as a coherent whole with them	The structured and unstructured data that are readily available between the organisation and the facilities are effectively managed and controlled through defined security and authorisation measures. Data is automatically available, but access must be approved at different locations before it can be accessed and used there	The methods and procedures to make data available between the organisation and the facilities are continuously improved to ensure data is securely available and accessible. The interoperability methods are continuously improved so that data is automatically exchanged and usable between the organisation and the facilities

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Data availability between the organisation and other providers (provider interoperability)	PA	Data is sent and received between the organisation and other providers to make relevant data available across different entities. Initial procedures for sending and receiving structured and unstructured data are ineffective and data is sent and received infrequently	Relevant data is frequently sent and received between the organisation and other providers to make relevant data available across different entities. Basic, repeatable sending and receiving methods and procedures are in place. Sent and received data are often in different formats and the systems are not interoperable. Basic procedures are established to make structured and unstructured data from and to different entities exchangeable and usable between them	Structured and unstructured data are readily and automatically available between the organisation and other providers. The effective interoperability of the organisation between other providers are established through standardised operating frameworks and protocols for the effective exchange and use of data between entities. The organisation interacts with other providers through local and wide-area networks to function as a coherent whole with them. Data between different entities are consistent	The structured and unstructured data that are readily available between the organisation and other providers are effectively managed and controlled through defined security and authorisation measures. Data is automatically available, but access to different relevant data must be approved for providers before they can access it	The methods and procedures to make data available between the organisation and other providers are continuously improved to ensure data is securely available and accessible. The interoperability methods are continuously improved so that data is automatically exchange and usability between the organisation and other providers
O	Technology for data sharing with the facilities and other care providers	SS	Technologies for data sharing from and to different facilities and other care providers exist. The available technologies gives very limited support to accomplish effective data sharing	Data sharing technologies are effective, robust, reliable and available. The proper hardware and software are available to repeatedly share data from and to different relevant entities	Data sharing technologies are compatible and interoperable with other devices and applications from different entities. Data sharing technologies support the effective functioning of data sharing activities as a coherent whole	Availability, reliability and maintainability of data sharing technologies are monitored and corrective maintenance is executed effectively and timely when technology failure occurs	Preventative maintenance and upgrades of data sharing technologies are executed timely and effectively to ensure that data sharing technologies continues to effectively support the sharing of data between relevant entities. Technologies for sharing data from structured and from unstructured repositories are continuously upgraded and updated

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Infrastructure for data sharing with the facilities and other care providers	SS	The data sharing infrastructure gives limited support to the sharing of data from and to the different relevant entities. Data sharing infrastructure is ineffective and unreliable	Data sharing infrastructure is available to consistently support the sharing of data from and to the different relevant entities	Specific data sharing infrastructure is applied to support the effective sharing of data from and to different relevant entities. The infrastructure support the sharing of data so that it can be used as a coherent whole through the different entities	The availability, reliability and maintainability of the data sharing infrastructure are measured and monitored. The data sharing infrastructure is repaired or replaced proactively when failure occurs	Data sharing infrastructure is continuously maintained and upgraded whenever needed to ensure data sharing infrastructure continues to support the seamless and effective sharing of data between different entities. Infrastructure is maintained or replaced before failure happens. Infrastructure is continuously improved to new innovative or alternative structures
Analysis							
F	Data aggregation for analysis	PA	Data aggregation for data analysis happens ad hoc and only when data is needed for analysis. Data aggregation happens manually and is resource-intensive. No defined procedure or methods are followed when data is aggregated for analysis	Data for analysis is aggregated on a consistent basis. Data is manually aggregated routinely following basic procedures and methods that enables the repeatability of aggregation. Data like claims and cost data, pharmaceutical and research and development data, clinical data, and patient behaviour are consistently aggregated	Data aggregation for analysis methods and procedures are effective and efficient. These defined methods and procedures are well-documented and followed to maintain the high level of effectiveness and efficiency	The methods and procedures of data aggregation for analysis are supervised and managed. The accurate and consistent execution of different data aggregation methods and procedures are monitored to detect when and where corrective action is needed	The effectiveness, efficiency, velocity and volume of data aggregation are continuously improved. More data that will contribute to better insights are continuously added to be aggregated for analysis. Aggregation is continuously improved towards automated aggregation of data for analysis

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Pre-processing for analysis	PA	Pre-processing happens when data is needed for analysis and on an ad hoc basis. Data pre-processing are carried out manually and is resource-intensive and slow. Data is not cleaned or reduced thoroughly and effectively. Data pre-processing is disorganised and no defined procedures or methods are followed	Pre-processing of analysis data happens on a consistent basis. Data is cleaned and reduced manually following basic procedures and methods that are carried out routinely. The raw data is consistently turned into usable data which can easily be interpreted	The data pre-processing procedures and methods are effective and efficient to clean and reduce raw data so that only usable data remains. The data pre-processing procedures and methods are well-documented. Data in different formats are transformed so that it can be stored and analysed. Pre-processing effectively decreases storage requirements and improves analytic accuracy	Pre-processing is stable and is executed on a predictable high level of effectiveness and efficiency. Pre-processing methods and procedures are monitored to detect when and where corrective action is needed	The effectiveness and the range of the ability of data integration, transformation, error detection and elimination, and data cleaning methods and procedures are continuously improved through exploiting pre-processing best practices innovatively. Pre-processing is continuously improved towards automated pre-processing
F	Storage for analysis	PA	Pre-processed and analysed data are stored using manual methods and procedures that are resource-intensive and is not carried out on a consistent basis	Basic methods and procedures for manual data storage are established that ensure that pre-processed and analysed data are consistently stored in an organised way. Structured data is stored consistently	Storage methods and procedures are effective and efficient and not resource-intensive. Data is stored in organised databases that are easily scalable. The storage space is easily accessible for the effective analysis of data	The efficiency and effectiveness of the data storage are monitored through well-established performance metrics. Inefficient data storage methods and procedures are identified and corrective action is taken proactively to ensure efficient and effective data storage	The data storage of analysis data is continuously improved to ensure pre-processed and analysed data is stored effectively and efficiently through exploiting data storage best practices. Storage is continuously improved towards automated storage

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Data analysis	PA	Descriptive analytics are incorporated. Analysis is resource-intensive and limited value can be gained from the analysis. No clearly defined methods, procedures, tools and techniques are incorporated. Analysis is done in silos and a limited variety of data types and formats are analysed	Diagnostic analytics are incorporated. Basic methods, procedures, tools and techniques are incorporated. A simple set of data types and formats that can be analysed are defined. Analysis is carried out routinely	Predictive data analytics are incorporated. Methods, procedures, tools and techniques are effective and well-documented. Integrated data analysis is carried out for comprehensive insights	Prescriptive data analytics may be incorporated. Data analysis is executed consistently and the effective execution of data analysis methods, procedures, tools and techniques are monitored and corrective action is taken. Data analysis can be carried out on many different types and formats of data.	The range of ability of data analysis is continuously improved through incorporating data analytics best practices. Tools and techniques are continuously updated to ensure the data analyse is optimised for use, efficiency and effectiveness. Data analysis is continuously improved towards automated analysis
F	Data analysis privacy and security	EP	Inadequate initial data analysis security and privacy checks and procedures are applied for the confidentiality, availability and integrity of data being analysed. The security and privacy of data being analysed cannot consistently be maintained	Basic data analysis privacy and security measures are applied for consistent security and privacy of data being analysed. Authorised access to devices, servers and databases used for the analysis of data is implemented to prevent unauthorised access and tampering of analysis data	Effective and rigorous data rules and control mechanisms for highly sensitive clinical data that is analysed are defined and applied to maintain the availability, confidentiality and integrity of patient data. Effective policies, standards, and compliance requirements to restrict the permissions of data analysts are applied	The adherence to data analysis privacy and security policies, standards and compliance requirements are monitored. The confidentiality and integrity of data being analysed are being monitored to ensure secure data analysis	Data analysis privacy and security policies, standards, and compliance requirements are continuously improved when necessary. The data analysis privacy and security tools and techniques are continuously improved to prevent of new threats. Potential threats are identified and measures are developed to address them

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Alignment to data analysis standards, policies and regulations	EP	Identified need for the alignment to data analysis standards, policies and regulations. No standards, policies or regulations are implemented or consistently aligned to	No alignment to data analysis standards, policies and regulations. If they are aligned to, it is incidental and unpredictable	The alignment to specific data analysis standards, policies and regulations ensures the effective functioning of data analysis. The standards, policies and regulations ensure that an organisation-wide understanding of activities, roles and responsibilities are in place	There are alignment to standards, policies and regulations control measures in place like auditing and alignment monitoring that allows for quantitative feedback	Complete alignment to data analysis standards, policies and regulations are carried out that enable strategy realisation. Any issues of non-alignment are identified and remedial action is taken to ensure alignment in a timely manner
F and O	Data analysis technology, tools and techniques	SS	Technologies for data aggregation, pre-processing, storage and analysis exist. The available technologies gives very limited support to accomplish effective data analysis	Data aggregation, pre-processing, storage and analysis technologies are adequate, robust, reliable and available. The proper hardware and software are available for consistent data analysis	Data aggregation, pre-processing, storage and analysis technology is compatible and interoperating with other devices and applications	Availability, reliability and maintainability of data aggregation, pre-processing, storage and analysis technology are monitored and corrective maintenance is executed effectively and timely when necessary	Preventative maintenance and upgrades of data aggregation, pre-processing, storage and analysis technology are executed timely and effectively. Technologies to analyse data are continuously upgraded and updated. New innovative technologies for analysis are continuously implemented
F and O	Data analysis infrastructure	SS	Existing infrastructure for data aggregation, pre-processing, storage and analysis gives very limited support to the effective execution of data analysis	Appropriate data aggregation, pre-processing, storage and analysis infrastructure is applied that gives adequate support for consistent data analysis	Specific infrastructure for data aggregation, pre-processing, storage and analysis is applied that supports data analysis effectively	The availability, reliability and maintainability of data aggregation, pre-processing, storage and analysis infrastructure are monitored. When failure occurs, infrastructure is repaired or replaced	Data aggregation, pre-processing, storage and analysis infrastructure is continuously maintained and upgraded whenever needed. Infrastructure is continuously improved to new innovative or alternative structures

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Security of data analysis devices	EP	Data analysis devices are not secured to enable the stable analysis of data. Analysis devices can easily be stolen that prevents effective analysis of data	Basic security policies and procedures are in place for the security of data analysis devices. Security policies and procedures does not completely ensure devices are secure and not stolen	Security policies and procedures for physical data analysis devices are well-defined and established that ensure a stable environment for data analysis through the devices	The security of physical data analysis devices are monitored. The correct usage of the devices are monitored to ensure it is not used for the wrong intent. Users that use the devices for the wrong reasons can be identified	The security of data analysis devices are continuously improved. Device security is continuously updated to prevent new threats and misuses of the devices
O	Data aggregation for analysis	PA	Structured and unstructured data aggregation for management and decision-making happens ad hoc and only when data is needed for analysis. Data aggregation happens manually resource-intensive and limited in the amount of data and variety of types and formats of data that can be aggregated. No defined procedure or methods are followed	The various structured and unstructured data from various sources are consistently aggregated. Data is manually aggregated routinely following basic methods that enables the repeatability of aggregation. Data can reliably and continuously be aggregated. Great amounts of different types and formats of data can be aggregated	Aggregation methods and procedures include tools and techniques that can effectively and efficiently aggregate extremely large amounts of various data types and formats from various sources. Data aggregation is fast and incorporates well-documented practices	The methods and procedures that enables the vast amount and variety of data aggregation are supervised and managed. The accurate and consistent execution of the different data aggregation methods and procedures are monitored to detect when and where corrective action is needed	The effectiveness and efficiency to aggregate great volumes and variety of structured and unstructured data at a high velocity is optimised. The range of ability and range of sources of aggregation is continuously improved. Data aggregation is continuously improved towards automated aggregation of data

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Pre-processing for analysis	PA	Pre-processing happens when data is needed for analysis and on an ad hoc basis. Data pre-processing are carried out manually and is resource-intensive and slow. Data is not cleaned or reduced thoroughly and effectively. Data pre-processing is disorganised and no defined procedures or methods are followed	Pre-processing of structured and unstructured data happens on a consistent basis. Data is cleaned and reduced manually following basic procedures and methods that are carried out routinely. The raw data is consistently turned into usable data which can easily be interpreted	The data pre-processing procedures and methods are effective and efficient to clean and reduce vast amounts and varieties of raw structured and unstructured data so that only usable data remains. The data pre-processing procedures and methods are well-documented. Data in different formats are transformed so that it can be stored and analysed. Pre-processing effectively decreases storage requirements and improves analytic accuracy	Pre-processing of vast amounts and varieties of structured and unstructured data is stable and is executed on a predictable high level of effectiveness and efficiency. Pre-processing methods and procedures are monitored to detect when and where corrective action is needed. Corrective action is proactively taken	The effectiveness and the range of the ability of data integration, transformation, error detection and elimination, and data cleaning methods and procedures to pre-process vast amounts and varieties of data are continuously improved. Pre-processed is optimised to decrease storage requirements, improve analytic accuracy and continuously improve automated pre-processing capabilities
O	Storage for analysis	PA	Pre-processed and analysed data are stored using manual methods and procedures that are resource-intensive and are not carried out on a consistent basis	Basic methods and procedures for manual data storage are established that ensure that pre-processed and analysed data are consistently stored in an organised way. Structured and unstructured data is stored consistently	Storage methods and procedures to store vast amounts and varieties of pre-processed and analysed data are effective and efficient and not resource-intensive. Data is stored in organised databases that are easily scalable. The storage space is organised and easily accessible for the effective analysis of data	The efficiency and effectiveness of data storage methods and procedures to store vast amounts and varieties of data are monitored through well-established performance metrics. Inefficient data storage methods and procedures are identified and corrective action is taken proactively to ensure efficient and effective data storage	The data storage of analysis data is continuously improved that allows greater volumes and varieties of pre-processed and analysed data to be stored effectively and efficiently through exploiting data storage best practices. Storage is continuously improved towards automated storage

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Data analysis	PA	Descriptive analytics are incorporated. Analysis is resource-intensive and limited value can be gained from organisation-wide analysis. No clearly defined methods, procedures, tools and techniques are incorporated. Analysis is done with a narrow scope and a limited variety of data types and formats are analysed	Diagnostic analytics are incorporated. Basic methods, procedures, tools and techniques are incorporated that are carried out routinely for organisation-wide analysis. A simple set of data types and formats that can be analysed are defined	Predictive data analytics are incorporated. Methods, procedures, tools and techniques can effectively analyse vast amounts and varieties of organisation-wide data and is well-documented. Integrated data analysis is carried out for comprehensive insights	Prescriptive data analytics are incorporated. Data analysis of organisation-wide data is executed consistently and the effective execution of data analysis methods, procedures, tools and techniques are monitored and corrective action is taken. Data analysis can be carried out on a vast amount and variety of different types and formats of data.	The range of ability to analyse vast amounts and varieties of organisation-wide structured and unstructured data is continuously improved towards automated analysis. New innovative ways to apply data analytics are exploited for better patient care, facility management, disease monitoring and organisation-wide decision making. Tools and techniques are continuously updated to ensure the data analyse is optimised for use, efficiency and effectiveness
O	Data analysis privacy and security	EP	Inadequate initial data analysis security and privacy checks and procedures are applied for the confidentiality, availability and integrity of data being analysed. The security and privacy of data being analysed cannot consistently be maintained	Basic data analysis privacy and security measures are applied for consistent security and privacy of data being analysed. Authorised access to devices, servers and databases used for the analysis of data is implemented to prevent unauthorised access and tampering of analysis data	Effective and rigorous data rules and control mechanisms for highly sensitive clinical data that is analysed are defined and applied to maintain the availability, confidentiality and integrity of patient data. Effective policies, standards, and compliance requirements to restrict the permissions of data analysts are applied.	The adherence to the specified data analysis privacy and security policies, standards, and compliance requirements for the analysis of vast amounts of data from various sources are monitored. The confidentiality and integrity of analysis data are being monitored to ensure secure data analysis	Privacy and security policies, standards, and compliance requirements for the analysis of vast amounts and variety of data from various sources are continuously improved when necessary. The data analysis privacy and security tools and techniques are continuously improved to prevent new threats. Potential threats are identified and measures are developed to address them

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Alignment to data analysis standards, policies and regulations	EP	Identified need for the alignment to data analysis standards, policies and regulations. No standards, policies or regulations are implemented or consistently aligned to	No alignment to data analysis standards, policies and regulations. If they are aligned to, it is incidental and unpredictable	The alignment to specific data analysis standards, policies and regulations ensures the effective functioning of data analysis. The standards, policies and regulations ensure that an organisation-wide understanding of activities, roles and responsibilities are in place. Ethical considerations of aggregating data from alternative sources for analysis are in alignment to the stated regulations	There are alignment to standards, policies and regulations control measures in place like auditing and alignment monitoring that allows for quantitative feedback. The regulations in terms of ethical considerations are monitored to ensure practices are aligned	Complete alignment to standards, policies and regulations with regards to the analysis of large amounts of data are carried out that enable strategy realisation. Any issues of non-alignment are identified and remedial action is taken to ensure alignment in a timely manner
					Usage		
F	Usage of shared patient data for integrated patient diagnosis and treatment	PA	Available data is used for patient diagnosis and treatment within the facility. Historic patient data is used for diagnosis and treatment. The use of data for diagnosis and treatment is sporadic and effective use of available data is dependent on individuals	Available patient data is used consistently in a disciplined way for patient diagnosis and treatment through the use of basic procedures and methods. Patient care data that is logged by different care professionals are used for consistent care	Procedures are effective to use available patient data logged by different care professionals for diagnosis and treatment through the comprehensive consideration of the data. Available data can be used to track patient medical history, interventions, encounters, lab tests results as well as managing allergies and drug contraindications for integrated patient care and treatment	The use of shared data for diagnosis and treatment is monitored to ensure the available data is used for its intended purpose. The care professionals understand the benefits of using the available data and of following the defined procedures for patient diagnosis and treatment and incorporated it into care delivery	The usage of shared data for integrated patient diagnosis and treatment is continuously improved. The procedures for the use of shared data is continuously improved for more effective diagnosis and treatment. The range of shared data use for diagnosis and treatment is continuously increased to aid in the diagnosis and treatment of more diseases and automatically recommend diagnosis and treatment

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Patient monitoring inside the facility	PA	Patient vitals are monitored in the conventional facility setting through using initial monitored methods. The use of data to monitor the patient data is limited. Procedures are not always followed and monitoring is often neglected	Data is consistently used to monitor patient vitals in the conventional facility setting with basic reliable and consistent methods	Methods and sensors for the effective monitoring of essential vital signs are applied for use in the conventional facility setting. Defined sensors such as electrocardiogram reading, heart rate, respiration rate, blood pressure, temperature, blood glucose levels and neural system activity are included. Effective procedures to execute monitoring are incorporated	The use of data for monitoring is controlled to ensure the data is used for the intended purpose of patient monitoring and that care professionals respond to monitoring indicators. The care professionals understand the benefits of using data to monitor patient vitals and procedures are rigorously followed	The use of data for patient monitoring is continuously improved through the use of monitoring best practices. The procedures for patient monitoring are continuously improved for the most effective patient monitoring. Monitoring is improved towards automated monitoring with indicators and notifications of changes
F	Patient monitoring outside the facility (remote monitoring)	PA	Patient vitals are monitored remotely through using initial monitored methods. The use of data to monitor the patient data is limited. The use of data for remote monitoring is expensive and not timely. Procedures and methods are not defined and followed seldom	Remote patients are monitored through the use of reliable communication and sensor methods. Remote patients ranging from chronically ill patients, elders, premature children to victims of accidents are monitored consistently through the continuous use of reliable data, following basic procedures and methods	Remote devices use data effectively for safety monitoring and adverse event prediction of patients. In-home monitoring is effective so that patients can carry on with normal activities without the hindrance of monitoring devices. The monitoring methods and procedures are well-documented and effective	The use of data for monitoring is controlled to ensure the data is used for the intended purpose of patient monitoring and that care professionals respond to monitoring indicators. The care professionals understand the benefits of using data to monitor patient vitals and procedures are rigorously followed	The use of data for remote safety monitoring and adverse event prediction is continuously improved through the use of monitoring best practices. The procedures for patient monitoring are continuously improved for the most effective patient monitoring so that patients can carry on with normal activities without hindrances. Monitoring is improved towards automated monitoring with indicators and notifications of changes

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Data usage for reporting on facility performance	PA	Data is used ad hoc and when necessary for facility performance. No clear procedures and KPIs for the reporting on facility performance exist. Data gives a limited indication of the performance of the facility	Data is used to consistently report KPIs of the performance of the facility. Data usage for the reporting of facility performance is carried out in a disciplined way through basic procedures and is done routinely. A basic set of KPIs are defined	The use of data for the reporting of facility performance is well-documented and effective. Specific meaningful KPIs are identified and data is used to effectively measure them for facility performance reporting	Data is used to measure the performance of the facility through defined KPIs. The data is used to manage and control the level facility performance of the KPIs through defined goals	The range of data use for reporting on facility performance is continuously enlarged to assist the facility in for example reducing errors, enhancing the acquisition of vital patient data, reducing needless expenditure, improving healthcare processes, doing preventative management, disease monitoring, diagnostic and therapeutic techniques and patient well-being
F	Utilisation of facility asset management	PA	Data is not used to track and determine the utilisation of facility assets	Data is used to track asset utilisation and to determine the availability of assets	Data is effectively used to track when assets are in use, for what period of time the asset is in use per session, the frequency of use and unnecessary breaks. The gap between the capacity and demand is effectively managed and bottlenecks are relieved	Data is used to schedule preventative maintenance at the right time to optimise asset utilisation. Data is used to determine anomalous or suspicious patterns in asset utilisation	The management of the utilisation of facility assets are continuously improved towards the automated management of assets through data. All assets are managed centrally and advanced asset management technology is successfully integrated into the workflows and normal responsibilities

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F	Medicinal stock level management	PA	Data is not used with defined methods and procedures to manage the medicinal stock levels. Stock is often not available when needed due to bad management	Data is used with basic methods and procedures to manage medicinal stock levels. Data is used to determine the demand for medicine at the facility	Effective methods and procedures are used to manage the medicinal stock levels. The facility knows the typical demand of all the different medicines for its locality and uses data to ensure medicine is available to treat the expected case load	The facility uses data to monitor its stock availability, stock demand and stock projections as it changes over time. Data is used to determine how much medicine to order, how long medicine last in stock and how long it takes to be delivered. Data is used to make forecasts	The facility continuously optimises its use of data to manage the medicinal stock levels. Medicinal buffer stock is optimised for maximum medicine availability and minimum storage cost. Data is used to determine stock levels in times of emergencies
F	Use of analysed data	PA	Analysed data is seldom used and no formal procedures are in place to ensure the use of analysed data. The use of analysis data has a limited application scope. Effective application of analysis data is sporadic and dependent on individuals	Analysed data is used consistently in a disciplined way following basic procedures that are established. This ensures that analysed data is consistently used for the determined scope of application	Analysed data is used effectively. Procedures are well-documented for the effective use of analysed data. The scope of application of analysed data is broad and analysed data is used as for decision making of the coherent whole. Decisions based on analysed data are applied proactively	The usage of analysed data is monitored. It is monitored whether analysed data is used for its intended purpose and whether the decisions made based on analysed data are applied effectively and immediately	The usage of analysed data is continuously improved. Data use procedures are continuously improved for more effective use of analysed data. The range of use of analysed data is continuously enlarged. Analysed data is used innovatively for better insights and decision making
F and O	Data queries for various usages	SS	Data query relations that exist give limited support to the usage of data. Limited queries are available to present stored data in different forms for different usages	Appropriate query relations exist to give adequate support for repeatable and consistent use of transactional data. Query relations support basic queries of stored data to represent it for basic usages	Data query relations that are established are specific for the effective usage of data and are well-defined. Query relations support comprehensive presentations of stored data for the effective views and usages of data from different data streams. Different data queries are established for different operational and governance purposes	The effectiveness of the query relations to support data usage are monitored. Ineffective query relations are addressed proactively to ensure that operational and governance purposes can be effectively executed through maintained data queries	Data query relations are continuously improved to ensure optimal data queries for data usage. Data queries are continuously updated to ensure the continuous effective support of new data usages

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Data usage technologies	SS	Technologies for data usage exist. The available technologies gives very limited support to accomplish effective data usage	Data usage technologies are adequate, robust, reliable and available. The proper hardware and software are available for repeatable data usage	Data usage technology is compatible and interoperating with other devices and applications. Technologies to use all available data are available	Availability, reliability and maintainability of data usage technology are monitored and corrective maintenance is executed effectively and timely when necessary	Preventative maintenance and upgrades of data usage technology are executed timely and effectively. Technologies to use all available data are continuously upgraded and updated
F and O	Data usage infrastructure	SS	Existing infrastructure for data usage gives very limited support to the effective execution of data usage	Appropriate data usage infrastructure is applied that gives adequate support for repeatable data usage	Specific infrastructure for data usage is applied that supports data usage effectively. Infrastructure for the use of all available are applied	The availability, reliability and maintainability of data usage infrastructure are monitored. When failure occurs infrastructure is repaired or replaced	Data usage infrastructure is continuously maintained and upgraded whenever needed. Infrastructure is continuously improved to new innovative or alternative structures
F and O	Data usage privacy and security	EP	Inadequate initial data usage security and privacy checks and procedures are applied for the confidentiality, availability and integrity of data being used. The security and privacy of data being used cannot consistently be maintained	Basic data usage privacy and security measures are applied for consistent security and privacy of data being used. Authorised access to devices, servers and databases used for the usage of data is implemented to prevent unauthorised access and usage of data	Effective and rigorous data rules and control mechanisms for highly sensitive clinical data that is analysed are defined and applied to maintain the availability, confidentiality and integrity of patient data. Effective policies, standards, and compliance requirements to restrict the permissions of data users are applied	The adherence to data usage privacy and security policies, standards and compliance requirements are monitored. The confidentiality and integrity of data being used are being monitored to ensure secure data usage	Data usage privacy and security software and control measures are continuously upgraded and updated to ensure the availability, integrity and confidentiality of data being used. The data usage privacy and security tools and techniques are continuously improved to prevent new threats. Potential threats are identified and measures are developed to address them

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
F and O	Security of data usage devices	EP	Data usage devices are not secured to enable the stable usage of data. Usage devices can easily be stolen that prevents effective usage of data	Basic security policies and procedures are in place for the security of physical data usage devices. Security policies and procedures does not completely ensure devices are secure and not stolen	Security policies and procedures for physical data usage devices are well-defined and established that ensure a stable environment for data usage through the devices	The security of physical data usage devices are monitored. The correct usage of the devices are monitored to ensure it is not used for the wrong intent. Policies and procedures are in place to monitor the security and use of data usage devices	The security of data usage devices are continuously improved. Device security is continuously updated to prevent new threats and misuses of the devices
F and O	Data usage align-ment to stan-dards, policies and regu-lations	EP	Identified need for the alignment to data usage standards, policies and regulations. No standards, policies or regulations are implemented or consistently aligned to	No alignment to data usage standards, policies and regulations. If they are aligned to, it is incidental and unpredictable	The alignment to specific data usage standards, policies and regulations ensures the effective functioning of data usage. The standards, policies and regulations ensure that an organisation-wide understanding of activities, roles and responsibilities are in place	There are alignment to standards, policies and regulations control measures in place like auditing and alignment monitoring that allows for quantitative feedback	Complete alignment to data usage standards, policies and regulations are carried out that enable strategy realisation. Any issues of non-alignment are identified and remedial action is taken to ensure alignment in a timely manner
O	Usage of data for coordi-nated care	PA	No clear methods and procedures exist to share patient data with all the relevant facilities and care providers to coordinate the delivery of effective care services	Basic procedures are in place to share patient data in order to coordinate care between different facilities and relevant care providers in a disciplined way. A basic set of data is used to coordinate care	Effective procedures and methods are established to share and use patient data as a coherent whole for the coordination of care between facilities like clinics, hospitals, homes, rehabilitation facilities, skilled nursing facilities and long-term care facilities and other care providers	The share and use of data are effectively incorporated into the coordination of care between different facilities and relevant care providers. Data is used to enhance teamwork across care settings, and is effectively shared and used for care management and medication management	The organisation continuously improves the coordination of care between different facilities. Best practices for the coordination of care are exploited to improve coordination of care and weak coordination of care practices are proactively strengthened

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Improve quality and efficiency of patient care	PA	Data is seldom used to improve quality and efficiency of patient care. No formal methods or procedures to use data to improve care exist	Basic procedures and methods exist to use data routinely to improve the quality and efficiency of patient care. The methods and procedures are carried out in a disciplined way	Specific data is defined and used to improve the quality and efficiency of patient care as a coherent whole. The methods and procedures to use data to improve the quality and efficiency of care are well-documented and effective. Data can help improve work practice standards, information availability, safety, quality of care and operational efficiency between facilities and administrative activities	The use of the defined data to improve the quality and efficiency of patient care is monitored to ensure that the data is utilised effectively. The execution of data usage methods and procedures are monitored to determine ineffective execution of data usage and where corrective action is needed	Data is used to continuously improve work practice standards, information availability, safety, quality of care and operational efficiency between facilities and administrative activities. The range of the use of data to improve the quality and efficiency of patient care is continuously enlarged
O	Usage of analysed data	PA	Organisation-wide analysed data is seldom used and no formal procedures are in place to ensure the use of analysed data. The use of analysis data has a limited application scope. Effective application of analysis data is sporadic and dependent on individuals	Organisation-wide analysed data is used consistently in a disciplined way following basic procedures that are established. This ensures that analysed data is consistently used for the determined scope of application	Large amounts of organisation-wide analysed data are used effectively for its determined purpose. The scope of applications for the use of analysed data is very broad and ranges applications like analysing disease patterns, tracking disease outbreaks and transmission to improve public health surveillance and speed response; faster development of more accurately targeted vaccines; and turning large amounts of data into actionable information that can be used to identify needs, provide services, and predict and prevent crises	The use of analysed data from big data analytics is supervised and monitored. It is monitored whether analysed data is used for its intended purpose and whether the decisions made based on analysed data are applied effectively and immediately	The range of applications of large amounts of organisation-wide data that are analysed are continuously introduced to improve data-driven decision-making, health delivery and spread of disease tracking for more effective healthcare delivery. The best practices of the analysis of large amounts of data are exploited innovatively to continuously improve the use of analysed data

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Data usage for organisation performance reporting	PA	Data is used ad hoc and when necessary for organisation performance. No clear procedures and KPIs for the reporting on organisation performance exist. Data gives a limited indication of the performance of the organisation	Data is used to consistently report KPIs of the performance of the organisation. Data usage for the reporting of organisation performance is carried out in a disciplined way through basic procedures and is done routinely. A basic set of KPIs are defined	The use of data for the reporting of the organisation performance is well-documented and effective. Specific meaningful KPIs are identified and data is used to effectively measure them for organisation performance reporting	Data is used to measure the performance of the organisation through defined KPIs. The data is used to manage and control the level facility performance of the KPIs through defined goals	The range of data use for reporting on the organisation performance is continuously enlarged to assist the organisation to reduce errors, reduce needless expenditure, improve healthcare processes, preventative management and disease monitoring. Data usage for the improvement of the organisation performance is continuously enhanced
O	Oversight of healthcare delivery system	PA	Data is seldom used to assist the oversight of healthcare delivery. No specific methods or procedures are in place to use data to oversee the healthcare delivery system. Data is used to assist the oversight of healthcare delivery on an ad hoc basis and only when necessary	Basic methods and procedures are in place to use data for routine oversight of the healthcare delivery system in a disciplined way. The delivery system can be frequently overseen through the use of data. Available data consistently represents the healthcare delivery system at determined intervals	The methods and procedures to use data for the oversight of the healthcare delivery system is well-documented and effectively executed. Data effectively represents the current state of the healthcare delivery system. The data represents the healthcare delivery system as a coherent whole	The data that is used for the oversight of the healthcare delivery system is controlled and when data deviates from reality, corrective action practices are in place to ensure minimal deviations. Data is effectively incorporated to manage, and make changes and alterations to the healthcare delivery system	The use of data is continuously improved to represent the whole delivery system near real time. Oversight is improved towards automated oversight with indicators and notifications of changes. The range of healthcare delivery system management and decision-making capabilities through the use of data is continuously improved

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Population health surveillance	PA	Data is seldom used for the surveillance of population health. No formal methods and procedures are in place to use data for the surveillance of the population health. Data is used on an ad hoc basis and when needed for population health surveillance	Basic methods and procedures are in place to use data for routine surveillance of the population health in a disciplined way. Basic population determinants are used for the surveillance of the population health	The methods and procedures to use data for the surveillance of the population health is well-documented and effective. Population health indicators, diseases, infections, pathogens, risk factors and other factors or determinants that provide an indication on the health status of the population are specified to be monitored for population health surveillance	The use of data for the surveillance of the population health is used for effective management and decision-making that effects the health of the population. The surveillance of population health through data enables data-driven decision-making that effects the population. Population health surveillance is used to disseminate health recommendation guidelines to the population	The range of population health surveillance through the use of data is continuously increased to include more meaningful population health indicators. Surveillance is improved towards automated surveillance with indicators and notifications of changes
O	Data usage for workforce and human capital management		Data is seldom used for workforce and human capital management. The organisation does not know the extent of its workforce and human capital	Basic methods and procedures are in place to manage the workforce and human capital. The organisation uses data to track the extent of its workforce and human capital and makes basic decisions based on that		Data is used to monitor the demand for workforce at different settings and manages the workforce according to the demand. The skills, knowledge and experience of the workforce is tracked and applied where it is needed most. The workforce meets the demand	Data is used to optimise the workforce in terms of its availability and cost. Data is used to optimise the skills, knowledge and experience available at the different facilities. Data is used to make forecast employment needs, identify workforce shortage needs
Governance							

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Data strategy	EP	An initial data strategy exists that gives some guidance to the data management activities of the organisation. The data strategy does not effectively and consistently control data management activities. The data strategy neglects the ethical considerations with regard to healthcare data	The data strategy gives repeatable outcomes. The data strategy gives guidance to the data management activities and consistently controls the data management activities. The data strategy has a narrow focus and basic healthcare data ethical considerations are included	A comprehensive data strategy is defined for all the data management activities across the different domain components that addresses healthcare data ethical issues comprehensively. The data strategy specifies specific goals for the different domain components to achieve the appropriate data strategy for healthcare data management. The data strategy has a broad focus and directs all data management domain components, models, policies, rules and standards of the data architecture	The outcomes of the domain components are measured for alignment with the data strategy to ensure the data strategy is achieved. The alignment of domain components to the data strategy ensure the effective execution of healthcare data management to achieve specified goals. The strategy to monitor the compliance to the ethical policies are also determined	The data strategy is continuously improved when necessary to ensure that it optimises healthcare delivery and patient care outcomes. The data strategy is continuously refined to enhance the capabilities of the management of data

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Master data management	EP	The management of master data is not in place to ensure data integrity throughout the whole healthcare data management system. The management of master data is not effective to ensure the uniformity, accuracy, stewardship, semantic consistency and accountability of the master data	Basic management of master data is in place to enable the integrity of data throughout the healthcare data management system. Data integrity is not maintained in all scenarios. The basic master data management system generally enables the uniformity, accuracy and semantic consistency of the master data	Master data management is well-established and documented to effectively maintain data integrity throughout the whole healthcare data management system. The uniformity, accuracy and semantic consistency of master data are effectively maintained. Master data is managed from one place to prevent data redundancy and unnecessary cleaning activities. Product and person master data are well-established and distinguished from each other to manage them differently according to their purpose	Both the person master data and product master data are effectively managed and controlled by data stewards. The master data is monitored by the data stewards. Data stewards effectively manage the integrity of data throughout the system, as well as maintain data uniformity, accuracy and semantic consistency. Accountability measures for the management of master data are established	The management of master data is continuously improved to enable the further improvement of data integrity, uniformity, accuracy and semantic consistency throughout the whole healthcare data management system
O	Meta-data management	EP	No clear common definitions and understanding of the characteristics the data elements exist	A basic metadata dictionary exists to define the characteristics of the data elements and to relate data from multiple sources	A comprehensive metadata dictionary exists to define all common data-element definitions and to relate all data from all sources. The metadata dictionary defines data elements / variables, their use in indicators, data collection methods, time period of data collection and analysis techniques used	Changes to the healthcare data management system is monitored and the changes are taken into account by modifying the metadata when necessary to ensure existing data, applications or processes are not disrupted when healthcare data management actors ask new questions or when new data sources are added	The metadata dictionary is continuously optimised the disparate needs of the technical, administrative and health user groups of the data management system

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Standards, policies and regulations for primary activities	EP	No clear standards, policies and regulations exist for the primary activities of the domain component. Standards, policies and regulations does not create a stable environment for the functioning of primary activities and does not control the functioning of the primary activities	Basic standards, policies and regulations exist that enable the controlled and consistent execution of the primary activities of the domain component. Basic standards exist with regard to ethical consideration of data management in healthcare	Well-defined standards, policies and regulations for all the primary activities of all domain component are in accordance to the determined data strategy to give guidance to the domain activities. Effective standards, policies and regulations to address ethical healthcare data management issues exist. The standards, policies and regulations are disseminated organisation-wide	Standards, policies and regulations that enable and promote the quantification and measurement of primary activities exist to ensure primary activities maintain the realisation of the data strategy	The standards, policies and regulations of the primary activities of the domain component enable the continuous improvement of the functioning of the primary activities
O	Standards, policies and regulations for support structures	EP	No clear standards, policies and regulations exist for the support structures of the domain component. Standards, policies and regulations does not create a stable environment for the functioning of support structures and does not control the functioning of the support structures	Basic standards, policies and regulations exist that enable the continuous functioning of the support structures of the domain component. Basic standards, policies and regulations exist that restrict the use of support structures that can create ethical healthcare data management issues	Standards, policies and regulations for the support structures of the domain component are well defined in accordance to the determined data strategy to give guidance to the implementation and functioning of support structures. Effective standards, policies and regulations to restrict the use of support structures that can create ethical healthcare data management issues exist. Standards, policies and regulations are disseminated organisation-wide	Standards, policies and regulations that enable and promote the quantification and measurement of the operation of support structures exist to ensure support structures maintain the effective support of primary activities to realise the data strategy	The standards, policies and regulations of the support structures of the domain component enable the continuous improvement of the functioning of the support structures

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Data life cycle management	EP	The organisation has no formal plan for data life cycle management. Deletion of data may cause data anomalies and data integrity issues	A basic plan exist for consistently archiving data, maintaining data warehouses, testing and delivering different application systems and deleting and disposing of data. A basic plan exist to deal with the different types of data along the whole data life cycle. Data life cycle management does not comprehensively account for all the different scenarios	Well-defined standards are used for archiving data, maintaining data warehouses, testing and delivering different application systems and anonymising and disposing of data that enable the effective management of the data life cycle. The life cycle of different types of data is specified to determine what data should be anonymised at what stages and what data must be kept. The anonymisation of data does not create data integrity issues	Data is controlled effectively to ensure that data anomalies do not occur. Measures are in place to ensure all data is managed according to the stated data life cycle standards. Data anomalies due to ineffective life cycle management can be identified and effectively dealt with	The standards for archiving data, maintaining data warehouses, testing and delivering different application systems and deleting and disposing of data are continuously updated to continuously improve the data life cycle management activities
O	Data security and privacy management	EP	Some initial data security and privacy rules and control mechanisms exist to prevent security breaches and protect privacy. Data security and privacy management does not prevent the use of data that can cause ethical issues	Basic data security and privacy rules and control mechanisms exist that consistently prevent security breaches and protect privacy. The data security and privacy measures creates a stable and secure environment for the management of data and prevents basic ethical issues	Data security and privacy rules and control mechanisms are well defined that reliably and effectively prevent security breaches and protect privacy. Effective security and privacy policies, standards and compliance requirements are defined that restrict the permission of users to ensure the proper use of data and prevent intricate data ethical issues. Data security controls are well-defined and executed for management security, operational security and physical security	Adherence to the security and privacy policies, standards and compliance requirements are supervised and control mechanisms are monitored and controlled. The data security and privacy standards and control measures enable the effective monitoring of the security and privacy of data. Management security, operational security and physical security are measured and controlled	The data security and privacy standards, rules and control mechanisms are continuously improved and updated for optimal data security and privacy and to prevent new security and privacy threats. Flaws in the defined security and privacy policies, standards and compliance requirements are changed whenever they are identified

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Adherence to national/ governmental healthcare data management acts/ regulations/ policies	EP	The standards, policies, regulations and activities of the organisation are established independently of national acts/ policies/ regulations and the organisation is not aware of all the relevant acts/ regulations/ policies. The organisation is not certain what data it has the right to share with other entities and it is not certain whether it adheres to regulatory environment policies	The standards, policies, regulations and activities of the organisation partially adheres to national acts/ regulations/ policies. The standards, policies, regulations and activities of the organisation are established on national acts/ regulations/ policies, but the acts/ regulations/ policies are not adequately followed. The organisation is aware of what data it has the right to share with other entities, tries to adhere to it and strives to prevent ethical issues with patient data	The standards, policies, regulations and all organisation activities are established on all relevant national acts/ regulations/ policies and adheres to their requirements. The organisation ensures that it effectively operates within the bounds of their rights on what data they are allowed to share with other entities. The organisation defines its data processes to effectively prevent ethical issues with healthcare data	Compliance to all relevant national acts/ regulations/ policies are monitored to ensure that all the established organisation standards, policies, regulations and activities are carried out in accordance to the national acts/ regulations/ policies. The organisation has measures in place to ensure it operates within its rights on what data it is allowed to share with other entities. The organisation monitors national/ governmental acts/ regulations/ policies for changes	Changes in the national acts/ regulations/ policies are continuously monitored. When changes occur, the standards, policies, regulations and activities are proactively and effectively updated to enable complete adherence to national acts/ regulations/ policies. Changes in the rights of the organisation to share data with other entities are continuously monitored and proactively adhered to. New ethical issues are continuously addressed

Continued on next page

Continued from previous page

SL	CA	CC	1. Initial	2. Repeatable	3. Defined	4. Managed	5. Optimising
O	Data business rules	EP	Initial business rules that exist are inadequate to give effective guidance to actions and constraints on data used by the organisation. The business rules gives very limited assistance to the organisation to achieve its goals, assert business structure and control or influence the behaviours of the business. The business rules are not aligned across the organisation	Basic business rules and logic exist to give partial guidance to actions and constraints on some data used by the organisation. The existing business rules partially assist the organisation to achieve its goals, assert business structure and control or influence the behaviours of the business. The business rules are not fully aligned across the organisation. Existing business rules are rigid and not flexible	Effective business rules express actions and constraints on all data used by the organisation. Data points are labelled and categorised to determine how they are dealt with. Business rules and logic are effectively applied on different data points to derive applications and inferences. Data business rules effectively help the organisation achieve its goals, assert business structure and control or influence the behaviours of the business activities. The business rules are aligned across the organisation and are implemented at all relevant locations like the code base, the CRM system, on the data processing and reporting side	The existing business rules are managed and regularly inspected for relevance. Ineffective or outdated business rules are retired and new relevant and effective business rules are introduced. The existing business rules are not rigid and can easily be altered or replaced. Business rules are effectively managed through data stewards	Business rules are and logic at the different locations like the code base, the CRM system, data processing and reporting side continuously improved to ensure the most effective assistance in achieving organisation goals, asserting business structure and to control or influence the behaviour of the business activities
O	Health indicators data collection specification	EP	The organisation has not specified any health indicators that its facilities must collect	The organisation has specified its own list of health indicators which is not necessarily aligned to global standards	The organisation has specified core health indicators which includes health status, risk factors, service coverage and health systems in accordance with global standards and specified effective collection policies, regulations and standards	The organisation monitors changes in the global reference list of health indicators and makes adjustments to its specified list when necessary	The organisation continuously improves its policies, regulations and standards to collect health indicators on the facility level and ensures that it remains aligned to global standards

Appendix F

Hypothetical results of the HCDMMM

The results sheet include an explanation of the presented results, along with indicators of how the results should be interpreted and a table with an overview of the results per domain component and capability category (Figure F.1). The results sheet also includes visual representations of the overall results (Figure F.2). Lastly, a table is included with the scores of all the capability areas which can be filtered according to different criteria in order to customise the user's view of the results (Figure F.3). A visual representation of the customised views is also available (Figure F.4). The results presented in Figures F.1 to F.4 are hypothetical results, and these figures only serve to present the HCDMMM.

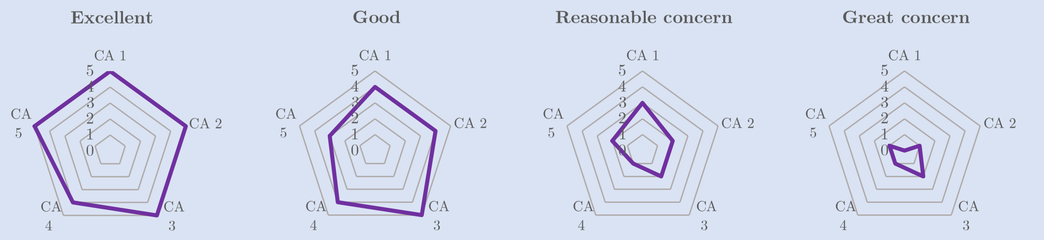
Results

Back to instructions | Back to organisational level | Clear all data

All relavent data can be viewed on this sheet for the organisation assessment. This sheet contains:
1) An overview of the average scores per domain component and per capability category
2) Visual representations of the overview data
3) Results of all capability areas in tabular and graphical format that can be customised

When viewing the results, note the following indicators:
-The result scores are highlighted according to a green-yellow-red colour scale
1) Green indicates the components that received the best scores of the assessment, which does not have to be focused on for improvement
2) Yellow indicates average scores. They are not the weakest areas, but are not functioning well enough, and might still be considered for improvement
3) Red indicates all areas that received the lowest scores that are the problem areas which need to be addressed. These are the focus areas for improvement

-The following radar charts give an indication of how to interpret the radar charts of the results:



Overview of results

	Primary activity (PAs)	Enabling practices (EPs)	Supporting structures (SS)	
Data collection	3,50	3,29	2,50	3,23
Data storage	3,00	4,50	2,00	3,40
Data sharing	2,67	2,33	3,33	2,78
Data analysis	2,50	4,67	4,00	3,56
Data usage	2,86	2,33	4,67	3,15
Data governance	-	3,13	-	3,13
	2,91	3,36	3,42	Overall

Figure F.1: Hypothetical results of organisational level assessment

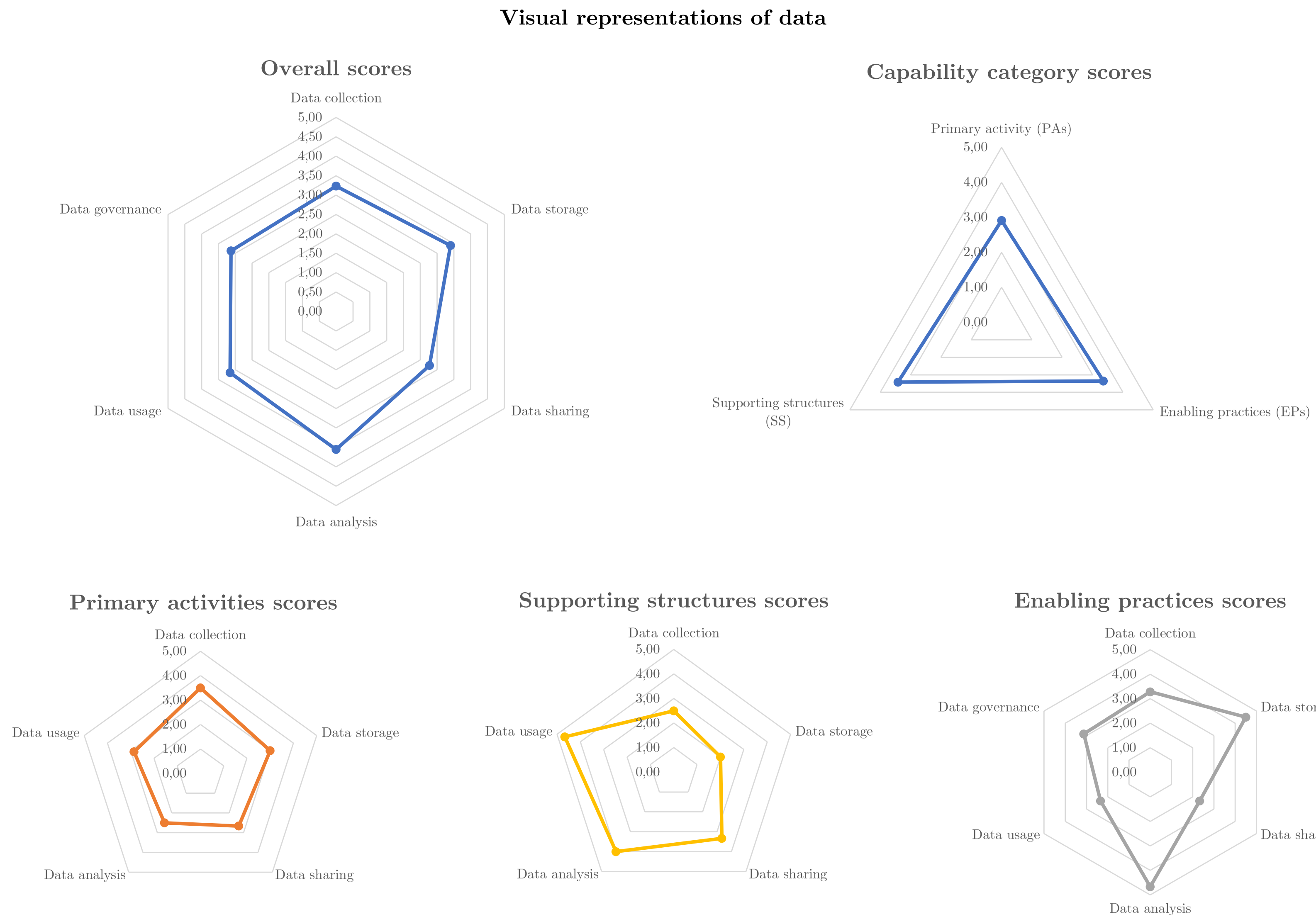


Figure F.2: Hypothetical results of organisational level assessment continued

Instructions for customised views of data

The table below can be used to create customised views of specific capability areas. The customised view will appear in the table below, as well as the graph above. Specific capability areas can be revisited on the assessment sheets by simply clicking on the capability area's name in the table.

Instructions for customised views:
1) Scores of capability areas can be seen per specific domain components or capability categories by using the filters of the columns **"Domain component"** or **"Capability category"**
2) A combination of the filters of the two columns can be used to see capability areas for specific domain components and capability categories
3) Number filters under the **"Score"** column can be used to see specific capability areas with scores:
- below / above average
- below / above a specified score
- equal to a specified score
- between specified scores

Capability areas results table

Domain component	Capability category	Capability area	Score
Data analysis	EP	Alignment to data analysis standards, policies and regulations	5
Data analysis	PA	Data aggregation for analysis	2
Data analysis	PA	Data analysis	3
Data analysis	SS	Data analysis infrastructure	4
Data analysis	EP	Data analysis privacy and security	5
Data analysis	SS	Data analysis technology, tools and techniques	4
Data analysis	PA	Pre-processing for analysis	2
Data analysis	EP	Security of data analysis devices	4
Data analysis	PA	Storage for analysis	3

Figure F.3: Hypothetical results of organisational level assessment continued

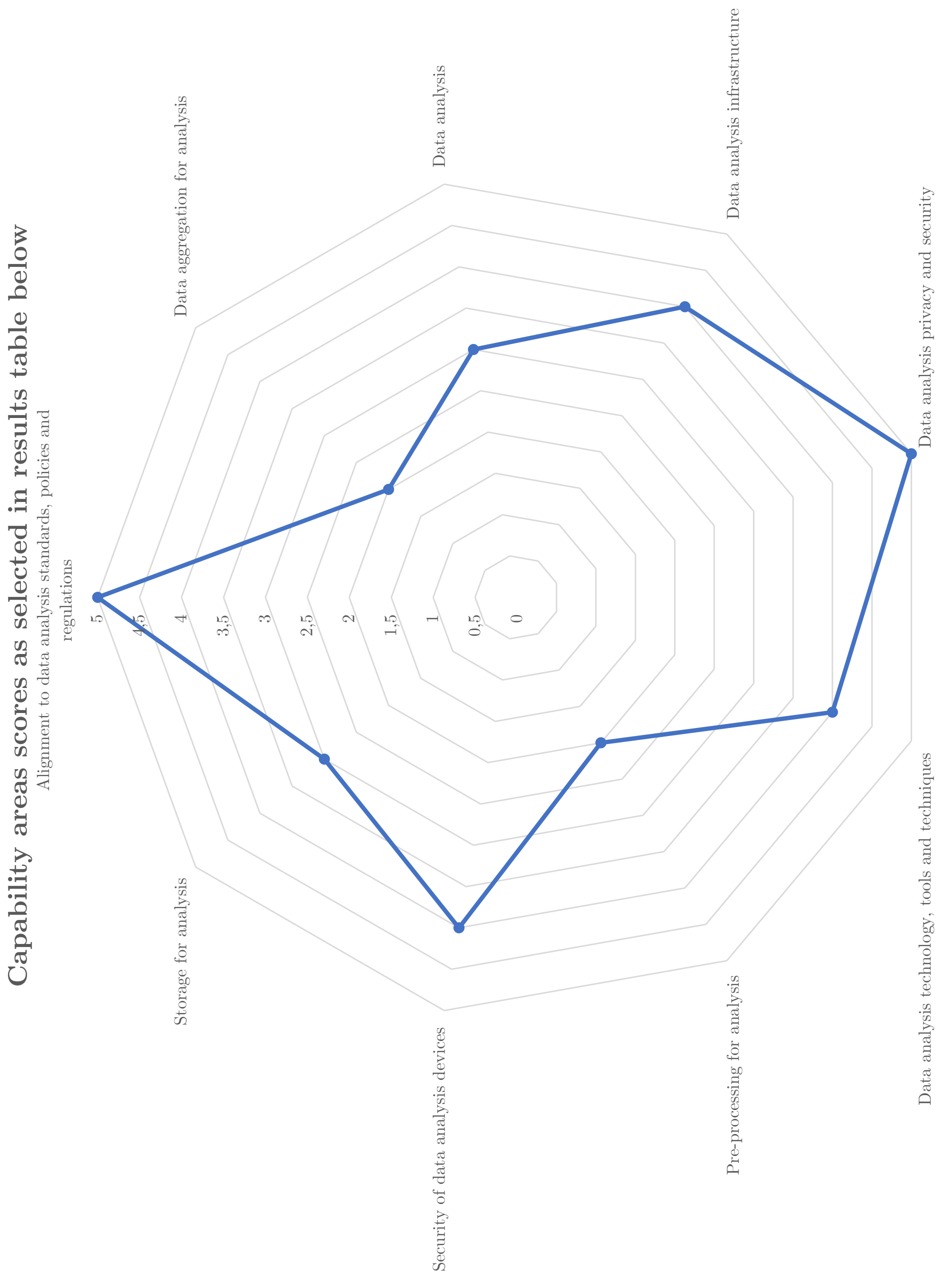


Figure F.4: Hypothetical results of organisational level assessment continued

Appendix G

Verification interviews summaries

The following sections summarise the interviews and correspondences with the different verification SMEs. All the SMEs were given a background of the study, a background on maturity models and the purpose of this maturity model. The different components of the model were also explained briefly. Different questions were presented to the different SMEs that fitted to their knowledge domains.

G.1 Maturity model SME - MS

The purpose of using a maturity model SME was to verify the logic of the maturity progression. The model included different maturation paths for different capability categories. The SME was used to verify the three different maturation paths and whether their progression is logical. This correspondence was not in the form of an interview. A document describing how the maturity levels of the different capability categories were determined, and what the definitions of their maturity levels are, were sent to the SME. He scrutinised the logic behind the different maturation paths and commented on the errors that he identified then. This correspondence was held after iteration five of the development process.

G.1.1 Summary of SME's comments

MS was a bit confused by the maturity progression of the capability areas. He was confused about the description of maturation that included both the process improvement of capability areas and capability areas progressing from paper-based to electronic. He was unsure whether the model describes a spectrum of ad-hoc to optimised maturity for paper-based data management, and then again from ad-hoc to optimised for electronic data management, or whether improving the maturity from ad-hoc to optimised have a direct impact on the maturation of a paper-based system into an electronic data management system.

MS agreed with most of the logic and the approach that was taken to specify the different maturation paths for different capability categories, the understanding of maturity levels and the determination of the maturity levels. However, the one concern that he highlighted was that the maturation of the capability areas tried to simultaneously describe the maturity progression as continuously improving processes and also describing the maturation from paper-based to electronic processes.

As an example to clarify what he means, MS pointed out that the maturity model does not allow for the description of an electronic system that has "level 2" maturity. He elaborated his example by stating that the fact that a capability area is electronic, does not automatically mean that it must be at least on maturity level 3. Another example he gave was that the "repeatability" (level 3) of a component can not automatically imply that the component is now thought of as electronic, rather than paper-based. One more example he gave was there can be a paper-based system that is "repeatable" that does not require electronic functionality, but at that time the maturation did not accommodate that.

Therefore, it was concluded that the progression of maturity levels that describes maturity from paper-based to electronic using the CMM-levels had flaws and needed to be changed.

Another important comment was that MS could not identify a significant difference between the maturation description for primary activities and supporting structures. Apparently all levels and names were the same. Level descriptors and definitions were essentially the same just with "primary activities" changed for "supporting activities".

G.1.2 Changes made to the model after the correspondence

- The definition of maturity concept was changed to process maturity (how well a capability is carried out), without referring to paper-based or electronic systems. This means that the description of maturation is not dependent on whether a capability area is paper-based or electronic, but rather maturity is dependent on how well that capability area is performed. This means the maturity model can assess both paper-based and electronic systems across the whole spectrum of the maturity levels identified.

- The number of maturity levels was changed to the five CMM-levels, which are *Initial*, *Repeatable*, *Defined*, *Managed* and *Optimising*.
- The maturity level definitions were changed which affected all maturity level descriptions of all capability areas across all the domain components. The maturation description of supporting structures were improved so that it is clearly distinct from the maturation of primary activities.

G.2 SQL and enterprise architecture SME - AV

The purpose of using an enterprise architecture SME with SQL understanding was to verify the basic structure of the maturity model with its different components in terms of whether the overall structure of the model comprise all the necessary enterprise and data structures. The SME was used to verify the domain components, capability areas and system levels included in the design architecture. This interview was held after iteration five of the development process.

G.2.1 Summary of discussion

The discussion comprised of two parts. The purpose of the first part of the discussion was to verify the overall architecture of the model and during the second part of the discussion specific parts were focused on.

G.2.1.1 The overall architecture

AV confirmed the viability of system levels and different model components. The structure and naming conventions are different to what she is used to, but has the same purpose.

AV confirmed that the system consists out of primary activities, support structures and enabling practices, but said that they used a framework that is different with other naming conventions.

AV confirmed the domain components that are included in the model as the necessary data management components.

AV asked for clarity on the data usage domain component. After explaining that it describes how the data is actually used for different purposes specifically for healthcare, she agreed.

AV questioned whether security and privacy management should be under the data governance domain component. She suggested that I check that again.

AV mentioned that the data structure should be changeable.

G.2.1.2 Specific components

Data capturing should mature to collecting basic patient data from home affairs if it is possible, but this is a very specific application of mature data collection.

AV questioned the function of "Data capturing for management and decision making" on the organisational level. She questioned why is it separately stated from "Data capturing" on the facility level. After explaining that data needed to be captured on the organisational level for other purposes than that of on the facility level, she agreed.

AV suggested adding data processing, but did not specify where. After some consideration, it was decided to consult literature further to determine where to include it.

The structure of data storage is important. She mentioned attribute dependency diagrams, but did not elaborate on it. She suggested looking at her company's report by Engelbrecht and Edwards for more on this. She further explained that the structure of the database can either enable or limit necessary functions.

AV suggested the inclusion of business rules in the model. She did not specify where. Mentioned rule engines versus hard coding to describe mature or immature business rules. She explained that business rules can be used to identify medicine that works or pandemic identification.

AV was very clear on the importance of the compliance to the Protection of Personal Information (POPI) Act for the confidentiality and security of data.

G.2.2 Changes made to the model after the interview

- Data processing was included under the data collection domain component as the transformation of raw data into usable and understandable information. Data pre-processing was already included under data analysis.

- After re-checking security and privacy management under data governance domain component, it was decided to keep unchanged as Yang *et al.* (2015) defined data security and privacy under data governance.

- Describing data capturing to mature to collecting basic patient data from home affairs is a very specific application of mature data collection. This model focus on high level, conceptual maturity descriptions and therefore, this description was not included under the data capturing capability area.
- The descriptions of data storage was not changed to describe data dependency diagrams, as they are outside of the scope of the study. It is part of the logical model design level and not the focus area which is the conceptual level.
- Business rules were included under the data governance domain component, because business rules assist in the effective management of data.
- The changeability of data structures were described under the model maturity level "Optimising".
- To include the adherence to the POPIA, a capability area for the adherence to all national acts, policies and regulations were included under the data governance domain component.

G.3 Data engineering SMEs

The purpose of the discussions with the data engineering SMEs were to verify that all the capability areas of the different domain components were developed right from a data engineering perspective. They assisted in verifying whether the data management capability areas were right and that their maturity descriptions are also theoretically sound. These discussion were held after iteration six of the development process.

G.3.1 Data engineer - MA

The purpose of the discussion with this data engineering SME was to verify the conceptual grounds of the data management architecture. Specific questions was also asked to address specific capability areas that has not been verified previously. This interview was held after iteration six of the development process.

G.3.1.1 Summary of discussion

MA mostly confirmed the conceptual grounds of the model and why it is an accurate representation of data management. The main contribution that MA made was to verify that the transactional information system and the big data system should be used complementary.

MA described big data as a methods to deal with the symptoms of existing data management systems that are functioning in silos. He described that it gets data in one place and in one format to be accessible and usable to all. MA said that big data treats the symptoms of different systems and makes collaboration between them possible.

MA confirmed the use of the information system that manages transactions. He stated that this component cannot be replaced by big data functions and processes. big data processes deal with differences between different systems and fix it for analysis and use.

MA suggested looking at more nuances of big data like ethical considerations.

MA is of the opinion that one should not be too specific to define processes as big data processes, but rather describe the capabilities.

MA confirmed that big data focuses on extracting value. He mentioned five V's (value, volume, velocity, variety and veracity) to define big Data.

MA stated that there are not that much value to be gained from in aggregating data from sources outside of healthcare system. He stated that it is messy, has ethical complications, and does not have much value to offer.

The optimising level should continuously look at additional functions that improve processes. For instance, additional sensors for data collecting.

MA also agreed with AV that processing of collected data is needed for transactional purposes. He elaborated that data processing should also be included at other domain components such as quality checks, processing for sharing, processing for storage, and processing for analysis (which was already included).

MA state that data sharing should be more explicit on what are authorised to be shared.

G.3.2 Changes made to the model after the interview

- The data storage and sharing capability areas were revised so that is does not distinguish between different storage systems for structured and unstructured data, but that it rather describe the capabilities of being able to store and share it without specifying how it is done.
- Ethical considerations of big data collection were included under the adherence to national or governmental acts, policies and regulations.

- Data processing for data sharing was included too, as data curation is present in all the domain components and not a separate domain component.
- The addition of functions at the optimising level of the capability areas was stated more explicitly.

G.3.3 Data engineer - GL

The purpose of the interview with this data engineer was to verify all the remaining parts of data management that has not been verified yet. Questions were focused on very specific components to gain certainty over these components. This interview was held after iteration six of the development process.

G.3.3.1 Summary of discussion

GL confirmed the capability areas that there was uncertain of under the data collection domain component: data entry forms, comprehensiveness of data captured, types of data captured and quality of collected data. He stated that the data collection capability areas should be extensive as it is currently.

GL confirmed data capturing for management and decision-making that AV questioned.

GL stated that it is necessary to include a capability area to validate captured data.

GL stated that data capturing should mature to automatic capturing, because there are many errors with manual capturing, but realised that it is not completely possible. He said that data capturing should be automated as far as possible, and then it should be validated.

GL stated that an objective of data management should be to eliminate paper-based capturing completely due to its errors and ineffectiveness, but realises that it is not completely possible.

GL suggested that data sending, receiving and interoperability should be merged and worded differently so that the capability area can mature towards not having to send and receive it, but that it is accessible and can be accessed through defined security and authorised measures. He suggested describing it as everything in the cloud and not confined to physical infrastructure. The idea is that sharing will happen automatically when this is allowed. Approval should be given before access at other facilities are enabled.

GL stated that data capturing and storage are easier when everything is in one system and not multiple systems.

GL is of the opinion that another objective should be to get rid of physical infrastructure for data management. Everything should move to the cloud. The security of the cloud is very good and it enables the efficient accessibility of data without the need to send or receive it. It is accessible everywhere, but security checks and authorisation should be in place to ensure it is not shared with the wrong users. Stated that cloud security is also less expensive. He suggested that the progression of data storing and sharing should be from physical infrastructure, to a hybrid system and then cloud-based.

GL suggested that the security of devices should be included across different domain components. The use of devices should also be monitored to ensure users use it appropriately.

GL agreed that the data usages on the facility level are focused more on medical purposes where the usages on the organisational level is focused on the managerial and generalised purposes.

GL confirmed the need of data queries at the data usage domain component, but said that the descriptions are currently too broad. He also mentioned two types of data queries: operational and reporting. Data queries are the representation of data in different forms to be able to use it for different usages. There are different facets to data queries: it can be on the CRM operational system or queries for reporting (data governance). Data streams should be defined to determine where data comes from and where does it go. Queries and logic that exist for data that is in a data warehouse, is applied to generate a table for the generation of a report to see what the state of a certain facility is.

GL said that although Master Data Management (MDM) is implicitly included in the model design, it should be explicitly included under data governance. MDM ensures data integrity throughout the system. MDM manages product data and person/customer (in this case healthcare) data throughout the system. Person master data is maintained by data stewards. Product master data relates to products and services (in healthcare it relates to medical products and services). Master Data (MD) is used across the whole system of the organisation. MD ensure proper data governance is in place and prevents data redundancy, because it is managed in one place and therefore, excessive cleaning processes are unnecessary. GL suggested that is should only be included under data governance, because it will become too technical to include it under every domain component.

GL stated that the data strategy is synonymous to the data architecture and that it changes continuously.

GL confirmed data life cycle as a data governance component. He suggested the inclusion of specification of policies for different data life cycle scenarios. Different types of data should be dealt with differently. Where in the life cycle the data is, should be considered. Different data should be removed or anonymised at a certain point and other data may be kept. He said that deleting data is not good for data integrity and that anonymising data maintains integrity. Policies for extreme cases should be included, for instance that a patient can ask for data or that it is removed.

GL also mentioned the importance of Acts like PoPI that regulates the sharing of data with other providers. The organisation should make sure it has the right to share the data.

GL confirmed the importance of business rules. He elaborated on the meaning of business rules. Business rules label data points according to different categories to determine how they are dealt with. There are business rules and logic applied on different data points to implement inferences. It also flags data points and categorises them to determine what happens to the data point. Business rules occur at four different locations: code base, the customer relationship management (CRM) system, at the data processing side, and the data reporting side.

GL mentioned training guides and training material so that personnel can know how to use the system. Facilitators ensure the correct use of the system.

G.3.3.2 Changes made to the model after the interview

- Validate captured data was included in the data collection domain component.
- Changed the description of data capturing to mature to automatic capturing as far as possible and the continuous addition of more data capturing processes that are automated.
- Data sending, data receiving and interoperability were merged and the maturation was changed that it starts from sending and receiving, but matures to one system where it is automatically available for access through secure channels.
- The capability area descriptions of physical infrastructure were changed to infrastructure so that the capability areas are not limited to infrastructure that is physical, but that it can allow the capability area to mature to not having to use physical infrastructure. Describing infrastructure in this way allows storage of data to move to the cloud.

- The security of devices and the monitoring of the correct use of these devices were included in the model.
- Two types of data queries were included (for reporting and monitoring) and the current descriptions were expanded.
- MDM was included under data governance to ensure the integrity of data throughout the whole system. The person and product master data were described and the role of data stewards.
- The data strategy capability area was elaborated to include the data architecture explicitly and to describe the continuous changing thereof.
- It was specified that different types of data should be dealt with differently in the life cycle management. It was specified that policies should be in place to delete or anonymise different data and that data of different scenarios should be dealt with differently.
- Stated more explicitly that adherence to Acts that describe the rights of the organisation to share certain data between facilities and different providers.
- Business rules were described better to describe how they label data points according to different categories to determine how they are dealt with. Business rules and logic are applied on different data points to implement inferences. It also flags data points and categorises them to determine what happens to the data point. Business rules occur at different locations: code base, CRM system, data processing side, the data reporting side.

G.4 Questionnaire correspondence with SMEs in the healthcare sector

The purpose of using healthcare data management SMEs was to determine how well the maturity model achieved the different requirements after the development process was completed. This was done by constructing a questionnaire with questions and a five-point scale that was used to determine how well different requirements were met through the model. Semi-structured interviews were also held with the SMEs to discuss further refinement of the model.

G.4.1 Healthcare data management SME 1

Healthcare data management SME 1 was a data scientist that was employed in the healthcare sector. He was included for the final verification phase as he had a comprehensive knowledge of data management in healthcare. This enabled him to verify all the components of the HCDMMM comprehensively.

G.4.1.1 SME 1 questionnaire answers

Figure G.1: Healthcare SME 1 questionnaire answers for verification

Questions or statements	Strongly agree (5)	Agree (4)	Unsure (3)	Disagree (2)	Strongly disagree (1)	Relevant model component
The use of the maturity model to assess an entity's health care data management leads to the eventual improvement of better care delivery to patients and better care delivery management and decision-making		x				Overall
Comment: I believe the maturity model can definitely be of tangible value in this regards, but due to the many factors that influence service delivery and management in public healthcare it will be extremely difficult to draw a direct a correlation between the use of this model and improved service delivery and management in a given healthcare entity.						
The maturity model can be used to make an as-is assessment of an entity's health care data management	x					Overall
Comment: This model can be very useful in benchmarking of a healthcare entity's health care data management. I have not come across a tool or model that creates this capability as well as this model does. Good combination of theoretical and practical.						
Do the maturity levels accumulate (improve incrementally), with each level encompassing the preceding lower level of maturity?		x				Overall
Comment: Some of the incremental levels are more exponential (in practice) that they are linear, therefore would require significantly more effort to move between maturity levels.						
To what extent do you agree with the capability areas included under data collection?		x				Data collection
Comment: One big capability that in my experience is an integral part of data collection is the capability of the human resources. Is the staff/data capturers trained, informed of practices and able to use the given technology if applicable? This may count as a type of supporting structure						
To what extent do you agree with the capability areas included under data storage?		x				Data storage
Comment: One aspect that I am somehow missing here is the ingestion of the collected data. In practice this a resource intensive process, that would be a significant contributing factor to the maturity of a given entity.						
To what extent do you agree with the capability areas included under data sharing?		x				Data sharing
Comment: One could also consider some enabling technologies as capabilities, as some sharing (especially externally through an API) requires specific technologies and capabilities.						
To what extent do you agree with the capability areas included under data analysis?		x				Data analysis
Comment: Would like to see the data analysis infrastructure fleshed out a little bit more. The definitions are a bit to generic to be able to judge in practice						
To what extent do you agree with the capability areas included under data usage?						Data usage
Comment: I am not sure if this includes all possible usage capabilities, but nothing additional comes to mind immediately						
To what extent do you agree with the capability areas included under data governance?		x				Data governance
Comment:						
To what extent do you agree with the two system levels, the facility- and organisational-level, that are included?	x					Overall
Comment: Depending on the use case, there might even be one higher system level, something like "governmental". Reason I say this is that the governmental data management maturity, trickles down to organisational level and subsequently to facility level. Yet I do not think it is necessary for this model to include this additional complexity and dependency, just worth the mention.						
To what extent do you agree with the supporting structures included in the model?		x				SS components

Comment: Mostly agree, but a few additional supporting structures might be needed here and there as mentioned in prior comments.						
The maturity model can be used by managers or change agents in health care entities to assess the maturity of the entity's data management		x				Overall
Comment: Definitions of the model just needs to be agreed upon by managers and or change agents, as some definitions are open for interpretation.						
The maturity model is user friendly and intuitive			x			Overall
Comment: Usage of this model will require some training, especially as it not something that is widely known. How will this model be assessed practically? Will it be done in this excel format or will a document/form be created.						
To what extent do you agree that the model is generic so that it can be used by more than one entity?		x				Overall
Comment: The healthcare sector is extremely diverse and therefore creating a generic model, that still allows for practically implementation is a challenge. I do believe this model can be used by more than one entity						
The maturity model uses standard domain language and is therefore, easily understandable			x			Overall
Comment: Not throughout, but domain language depends on if you speak to management, healthcare professionals or technical staff. I cannot expand in detail, as this would most likely be an essay in itself.						
The maturity model considers all the necessary governmental and national policies, acts and regulations			x			Under data governance
Comment: I am not qualified to comment on this						
The maturity model considers ethical considerations about data management in health care		x				Specifically, under data analysis on the organisational level: Alignment to data analysis standards, policies and regulations
Comment: It is included yes, but I am not convinced that the ethical considerations come through strongly enough in the model.						
The maturity model includes the privacy and security of data		x				Some EP components of all domain components
Comment: Except for one prior comment it accounts for privacy and security. Just a note: Privacy and security in the public healthcare sector is a hotly debated and contentious topic. For instance, it is still debated to today whether cloud storage or onsite storage is more secure and whether physical files are better than electronic files – both these have trivial answers, yet they remain points of contention.						
The maturity model includes all the necessary internal standards and policies that an entity should consider		x				Some EP components of all domain components
Comment: Internal standards might be highly specific to a given entity, but to my knowledge the model has nothing missing that is apparent						
The maturity model is not applicable to other domains other than health care data management					x	Overall
Comment: This model is specifically very applicable to the healthcare. Additionally, I believe that the baseline of this model could be applied to other data management domains. Adjustments would obviously be required in sections like Data Usage where the model is very specific to healthcare, but the premise of the model would still hold across domains.						
The maturity model does not specify any specific technology (like specific software) and infrastructure that should be incorporated to achieve a maturity level	x					Some SS components
Comment: Could terms like structured database or cloud computing still be used as point of reference? Or is this model supposed to age with time as well?						
The maturity model describes different maturity levels of capability areas without prescribing how to achieve maturity levels	x					Overall
Comment: I agree with this as road to maturity would be highly dependent on the given entity						

G.4.1.2 Changes made after correspondence

- Added data ingestion capability area
- Included ethical considerations under data governance more explicitly
- Developed the transfer media with introduction, overviews of system levels, instructions for the assessment methodology, domain component assessment sheets and results sheets so that the HCDMMM is user-friendly and practicable

G.4.2 Healthcare data management SME 2

Healthcare data management SME 2 was a process improvement engineer that was employed at a private hospital group in South Africa. He was included for the final verification phase as he had a comprehensive knowledge of data management in healthcare, and also had experience with maturity models. This enabled him to verify all the components of the HCDMMM comprehensively.

G.4.2.1 SME 2 questionnaire answers

Figure G.2: Healthcare SME 2 questionnaire answers for verification

Questions or statements	Strongly agree (5)	Agree (4)	Unsure (3)	Disagree (2)	Strongly disagree (1)	Relevant model component
The use of the maturity model to assess an entity's health care data management leads to the eventual improvement of better care delivery to patients and better care delivery management and decision-making		X				Overall
Comment:						
The maturity model can be used to make an as-is assessment of an entity's health care data management			X			Overall
Comment: Unsure about the practicality of the maturity model? Please apply to one hospital / group and validate the appropriateness of the questions? Also, different answers will be provided by different people... How will the model calibrate / deal with this? Too much room left for interpretation by assessor. Be careful for words that are unscientific and open for interpretation e.g. "adequate", "available", "reliable" etc. It seems that “maturity” is in the eye of the beholder and not scientific enough						
Do the maturity levels accumulate (improve incrementally), with each level encompassing the preceding lower level of maturity?				X		Overall
Comment: Due to vague terminology (see comment above) the lines become blurry between different maturity levels. When you try to apply this in practice, the employee who is responsible for data management will provide countless reasons why the current reality can be interpreted as “mature” and then it is “his word against yours”. The way around this is by asking questions with incontestable yes/no answers which add up to a score. A higher score resembles a higher maturity.						
To what extent do you agree with the capability areas included under data collection?		X				Data collection
Comment:						
To what extent do you agree with the capability areas included under data storage?		X				Data storage
Comment:						
To what extent do you agree with the capability areas included under data sharing?		X				Data sharing
Comment:						
To what extent do you agree with the capability areas included under data analysis?		X				Data analysis
Comment:						

To what extent do you agree with the capability areas included under data usage?		X				Data usage
Comment:						
To what extent do you agree with the capability areas included under data governance?		X				Data governance
Comment:						
To what extent do you agree with the two system levels, the facility- and organisational-level, that are included?		X				Overall
Comment:						
To what extent do you agree with the supporting structures included in the model?		X				SS components
Comment:						
The maturity model can be used by managers or change agents in health care entities to assess the maturity of the entity's data management				X		Overall
Comment: I question the practicality of the model.						
The maturity model is user friendly and intuitive					X	Overall
Comment: Refer to the comment in red above						
To what extent do you agree that the model is generic so that it can be used by more than one entity?		X				Overall
Comment: It is almost too generic.						
The maturity model uses standard domain language and is therefore, easily understandable			X			Overall
Comment: I'm not an expert in the field, but I think there will be vastly different interpretations of the maturity levels provided. Although standard "domain language" is provided, it is not concise enough (in my opinion)						
The maturity model considers all the necessary governmental and national policies, acts and regulations					X	Under data governance
Comment: I did not see any references made to specific policies, acts or regulations.						
The maturity model considers ethical considerations about data management in health care		X				Specifically, under data analysis on the organisational level: Alignment to data analysis standards, policies and regulations
Comment:						
The maturity model includes the privacy and security of data		X				Some EP components of all domain components

The maturity model includes all the necessary internal standards and policies that an entity should consider				X		Some EP components of all domain components
Comment: “Standards and policies” are mentioned but no reference made to specific known standards or definitions provided. This will complicate the application of the model.						
The maturity model is not applicable to other domains other than health care data management				X		Overall
Comment:						
The maturity model does not specify any specific technology (like specific software) and infrastructure that should be incorporated to achieve a maturity level	X					Some SS components
Comment:						
The maturity model describes different maturity levels of capability areas without prescribing how to achieve maturity levels	X					Overall
Comment: I strongly agree, but I’m unsure whether this is a strength or a weakness. When an organisation scores low on the maturity model it would want to know how to improve. The model is unfortunately too vague to assist in improving further (my opinion)						

G.4.2.2 Changes made after correspondence

- Improved maturity levels statements to be more quantitative and concise so that as-is assessments can be made
- Included capability areas of financial data capturing and storage for patient care
- Included capability areas of workforce management and human capital data capturing and storage
- Developed the transfer media with introduction, overviews of system levels, instructions for the assessment methodology, domain component assessment sheets and results sheets so that the HCDMMM is user-friendly and practicable
- Eliminated or clarified vague, unspecific or ambiguous terminology so that all users have a common understanding of the meaning

Appendix H

Requirements specification evaluation

H.1 The functional requirements verification

The specified functional requirements ensured that the core performance demands that the HCDMMM should meet, were developed. Table H.1 describes that the functional requirements that state the core performance demands and that ensure the problem is met comprehensively, were met.

Table H.1: Verification of the functional requirements

ID	**Requirement**	**Satisfaction description**	**Satisfaction components**
FR1	The framework should enable the improvement of healthcare data management that improves healthcare delivery to patients and improve care delivery management and decision-making	The HCDMMM assesses a healthcare entity's data management that leads to the eventual of better care delivery. It includes different capability areas with their different maturity levels to illustrate the different capabilities that a healthcare data management entity should have for care delivery and the maturity levels give an indication of the consecutive improvement steps of data management towards improved care delivery. The HCDMMM can assist the improvement of healthcare delivery when applied to assess the as-is state and to obtain an indication of the next maturity step	The overall HCDMMM assessment tool with its included capability areas and their maturity levels can be used to make a maturity assessment with the purpose of improving data management for better care delivery

Continued on next page

Continued from previous page

ID	**Requirement**	**Satisfaction description**	**Satisfaction components**
FR2	The framework should be able to be used to give an as-is assessment of healthcare data management	The HCDMMM has progressive, incremental, concise, specific and distinct maturity levels to make maturity assessments. The maturity model consists of different domain components and different capability areas. Each capability area has a set of progressive maturity levels with descriptions that enables the identification of the as-is state of the specific capability area. The instructions on how to use the HCDMMM assessment tool is straightforward and makes it intuitive to make capability maturity assessments. All the capability areas together gives an as-is state assessment of the healthcare data management system	The progressive, incremental, concise, specific and distinct maturity levels of the HCDMMM. The assessment instructions and incontestable method of assessment
FR3	The framework should describe the incremental improvement of healthcare data management	The HCDMMM has progressive, incremental, concise and distinct maturity levels. Each consecutive maturity level of the HCDMMM is built on its preceding level. The maturity model was design with a number of cumulative stages where higher stages built on the requirements of lower stages	The progressive maturation of maturity levels of all capability areas of the HCDMMM
FR4	The framework should consider all the important data management components relevant to healthcare	The HCDMMM was developed to describe all the different components of healthcare data management with their capability areas that ensures all the different aspects are included that enables the whole data management system	All included domain components with their capability areas

Continued on next page

Continued from previous page

ID	Requirement	Satisfaction description	Satisfaction components
FR5	The framework should include necessary system components to describe healthcare data management as a functioning system	The HCDMMM describes healthcare data management as a system with two different system levels that indicates that different functions are carried out on different system levels. The HCDMMM also describes different capability categories which represent different types of capability functions that work together as multiple sub-systems to achieve the goal of healthcare data management	The organisational and facility levels, the different capability categories and their interactions
FR6	The proposed research product should include the necessary technological and infrastructural components needed for healthcare data management	The HCDMMM describes the technological and infrastructural capabilities and their levels specifically for each domain component. The capabilities of the different technological and infrastructural components are described without specifying specific technologies or infrastructure	The technological and infrastructural capability areas of the domain components

H.2 The user requirements verification

The user requirements were determined from the viewpoint of the user and ensured the HCDMMM was developed for the intended user and that it is usable by that user. Table H.2 indicates that the user requirements were met by the HCDMMM.

Table H.2: Verification of the user requirements

ID	Requirement	Satisfaction description	Satisfaction components
UR1	The framework should be usable to managers or change agents of healthcare entities to enable data management improvement on a strategic level	The HCDMMM was developed from the view of strategic and managerial users. The components that were included in the HCDMMM were described to be of use to users that makes strategic and managerial public healthcare data management decisions. The HCDMMM allows the maturity assessment of healthcare data management on the facility and organisational level from the strategic and managerial viewpoint. The HCDMMM assessment tool includes introductory information that gives adequate information on the use of the model to managers. Instructions on how to make an assessment give direct and clear instructions on how users can make an incontestable maturity assessment. The method of assessment of the HCDMMM assessment tool is intuitive and easy to carry out	The HCDMMM was developed for a managerial and strategic audience. The HCDMMM assessment tool with background information, instructions and an intuitive assessment method makes the HCDMMM usable to managers and change agents
UR2	The framework should be generic so that it is usable to different healthcare entities	The HCDMMM was not developed based on the public healthcare data management structure of a specific country. All the generic components that public healthcare data management systems should include, were incorporated	The generic nature of the HCDMMM

Continued on next page

Continued from previous page

ID	**Requirement**	**Satisfaction description**	**Satisfaction components**
UR3	The framework should be user-friendly and intuitive	The HCDMMM was developed so that it is easily understood by the user. Background information and instructions on how to use the model to assess healthcare data management are given. Concept definitions and terminology clarification are incorporated to eliminate ambiguity and ensure a common understanding of terminology. The instructions gives an easy to follow assessment method and the navigation between sheets ensure user-friendliness. The maturity level descriptions are concise to ensure user-friendliness	Additional information that explains the model, clear and concise maturity level descriptions, defined concepts and clarification of terminology
UR4	The framework should use standard domain language to be easily understandable	The common terms of healthcare data management was used to describe the different components and maturity levels of the model to ensure it is easily understandable by users on a management and strategic level. Definitions of different components are given to eradicate any ambiguity and to clarify what is meant by different terminology used	Common data management and healthcare terminology for users on management and strategic level were used throughout the model. Domain concepts were defined and unspecific terminology were clarified

H.3 The boundary conditions verification

Boundary conditions should be met unconditionally. In Table H.3 it is described that the HCDMMM met the boundary conditions that ensured all the necessary legal and ethical considerations that an healthcare entity should comply to, were considered.

Table H.3: Verification of the boundary conditions

ID	**Requirement**	**Satisfaction description**	**Satisfaction components**
BC1	The proposed research product should consider governmental and national policies, acts and regulations	The HCDMMM incorporates the adherence to national and governmental policies, acts and regulations under the data management domain component. These national and governmental policies, acts and regulations also influences the other domain components to ensure all data management practices adhere to these policies, acts and regulations. These acts, policies and regulations are included generically without specifying and specific acts, policies and regulations of any countries to maintain the HCDMMM generic nature	Adherence to national/ governmental healthcare data management acts/ regulations/ policies, and alignment to standards, policies and regulations capability areas

Continued on next page

Continued from previous page

ID	Requirement	Satisfaction description	Satisfaction components
BC2	The proposed research product should describe the ethical considerations around healthcare data management	The HCDMMM describes that the policies should include ethical considerations with the management of internal and external health data that are aggregated in the healthcare data management system. This was also explicitly stated under the data analysis domain component. Most of the data governance domain components include ethical considerations in their maturity levels	Data governance capability areas, adherence to all domain component acts and policies, and specifically alignment to standards and policies of organisational data analysis
BC3	The proposed research product should incorporate the privacy and security of data	The HCDMMM describes the data security and privacy management under the data governance domain to ensure the security and privacy of healthcare data are maintained. Under all the other domain components data security and privacy are also described to ensure data privacy and security are incorporate across all the healthcare data management components. Many domain components include the security of devices used too	Privacy and security management of data governance, domain component capability areas that focus on data privacy and security and the security of data management devices

Continued on next page

Continued from previous page

ID	Requirement	Satisfaction description	Satisfaction components
BC4	The framework should include the necessary internal standards and policies of the entity under study	The HCDMMM describes the necessary internal standards, policies and regulations of primary activities and supporting structures that adhere to national and governmental acts, policies and regulations. Alignment to these standards, policies and regulations are also included under each of the domain components to ensure the entire healthcare data management system is aligned to the specified policies, standards and regulations	Data governance capability areas, enabling practices of domain components that ensure alignment to policies and standards

H.4 The design restrictions verification

That the HCDMMM met the design restrictions are described in Table H.4. The design restrictions determine the preferred solution space. All restrictions, limits and exclusions placed on the HCDMMM were verified.

Table H.4: Verification of the design restrictions

ID	Requirement	Satisfaction description	Satisfaction components
DR1	The framework is limited to only data management in the healthcare domain	The HCDMMM was developed with the specific focus of data management in healthcare and not all components are relevant to data management in general. Aspects that are essential to healthcare were included in the model which are not necessarily important to data management in other contexts. With slight alterations, the HCDMMM can be applied to other data management domains	The applicability of the HCDMMM in the healthcare data management domain
DR2	The proposed research product should not specify specific technologies, processes and methods to achieve its goal. The framework does not elaborate on the technical detail of healthcare data management systems	The capabilities of the different technologies, processes and methods are described without specifying specific technologies, processes and methods	Technology and infrastructure capability areas of all domain component

Continued on next page

Continued from previous page

ID	Requirement	Satisfaction description	Satisfaction components
DR3	The framework should describe different stages of healthcare data management improvement without prescribing how to achieve the improvements	The HCDMMM was developed as a descriptive maturity model that describes the different maturity levels of the capability areas, but it does not prescribe how to reach the maturity levels and what tools to use. The HCDMMM gives level descriptions of what the capability area should be able to accomplish at every maturity level. This enables the description of different improvement stages without prescribing how to achieve the improvements	The descriptive nature of the HCDMMM

H.5 The attention points verification

The attention points are relevant requirements that are desirable and should be noted and considered, but does not have to be met as a dictate. That the attention points were met by the limits and exclusions placed on the HCDMMM are conveyed in Table H.5.

Table H.5: Verification of the attention points

ID	Requirement	Satisfaction description	Satisfaction components
AP1	The desired focus of the framework is to develop it to assist the improvement of healthcare data management in developing countries, but does not have to be limited to developing countries	The scope of challenges and challenge landscape that are based on the challenges of developing countries that was considered during the design of the model caused the model to be applicable to healthcare data management of developing countries than developed countries. Developed countries can still benefit from the model if their healthcare data management structures are similar to the model's structure	Scope of model on developing countries
AP2	The proposed research product should strive towards the standardisation of all data management components across different units	The HCDMMM was developed with the aim to describe a maturity level where all components are standardised, with the addition of the processing and analysis of unstructured data to enable the analysis of large quantities of data in various formats as it will be impossible to standardise the whole system	Defined maturity level
AP3	It should be considered that best practices are evolving	Because best practices are evolving and does not remain the same, no best practice were included explicitly in the HCDMMM, but the optimising maturity level includes the consideration of current best practices	Optimising maturity level

Appendix I

Validation supporting content

This appendix contains the supporting content of the validation process as described in Chapter 6. The content is as follows:

- Section I.1: Validation questionnaire
- Section I.2: Validation SMEs' responses
  - Section I.2.1: Validation SME 1 responses
  - Section I.2.2: Validation SME 2 responses
  - Section I.2.3: Validation SME 3 responses
  - Section I.2.4: Validation SME 4 responses
  - Section I.2.5: Validation SME 5 responses
  - Section I.2.6: Validation SME 6 responses
  - Section I.2.7: Validation SME 7 responses

I.1 Validation questionnaire

This section contains the validation questionnaire that the validation SMEs completed to validate the developed model.

Figure I.1: Validation questionnaire

Validation of the Healthcare Data Management Maturity Model (HCDMMM)

16/09/2020

Participant information

Participant name:

Occupation/background relevant to this study:

The study aim:

The aim of the study was to develop a tool that enables the identification of the strengths and weaknesses of existing data management practices that assists systematic improvement initiatives and planning.

Aim of the validation process:

The HCDMMM was developed to assist managers of health care entities to assess the maturity of their data management. The HCDMMM includes two system levels on which maturity assessment can be conducted on, namely the facility and the organisational level. On the facility level, clinics and hospitals can use the HCDMMM to make a data management maturity assessment and on the organisational level, the organisation's headquarters can use the HCDMMM to make a maturity assessment of their data management. To validate that the HCDMMM achieves this goal, it was assessed along the dimensions of:

- applicability to real-world cases: the conceptual structure and components of the HCDMMM are representative of the data management of real-world health care entities on a facility and an organisational level;
- practicability: the HCDMMM and its assessment methodology can be put into practice; and
- usability: the degree to which the HCDMMM can be used by the intended user.

Explanation of the questionnaire questions

The questionnaire is constructed to validate the HCDMMM according to the abovementioned dimensions. The focus of the study was that the HCDMMM should be applicable to the public sector. Therefore, the questionnaire includes questions that validate that the HCDMMM is applicable to the facility and organisational levels of the public health sector. Questions are also included to validate whether the HCDMMM is applicable to the private sector on these levels.

The questions seek to validate that the assessment method of the HCDMMM can be put into practice to make a maturity assessment of a health care entity. As the assessment method is the same for all levels and sectors, the practicality of the assessment method is only asked once in the questionnaire. The same reasoning applies to why the other validation dimension

Figure I.2: Validation questionnaire *continued*

questions (related to usability, uniqueness, and robustness) also do not specify the system levels or sectors.

Validation process

The validation process comprises of two stages. The first stage entails a presentation by the researcher explaining the artefact and the validation process, and the second entails the participants' input and feedback by means of completing the validation questionnaire. Each of these stages are elaborated on below:

1. Presentation:

All the participants of the validation process will be taken through the HCDMMM. The different sheets of the Excel model will be explained, including the landing page, the system level overviews, instructions, domain components assessment sheets and the results sheets. Navigation between sheets will be demonstrated and the assessment instructions will be explained. The validation process will also be explained. After the presentation, a discussion will follow to clarify any remaining questions about anything pertaining to the workings of the HCDMMM and the validation process.

2. Participant input / feedback:

After the presentation, participants are given the opportunity to complete the form below. For each question or statement with a scale, the participants mark an 'x' how strongly they agree with the statement or question. Where applicable and/or necessary, free-text space is provided for comments.

Should the participants have any additional comments, want to elaborate on their allocated scores or if they believe something is important to point out for the validation assessment, these may be provided below the each question in the table in the space provided.

Figure I.3: Validation questionnaire *continued*

Validation questionnaire

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
			5	4	3	2	1
1.1	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the public health sector					
Comment:							
1.2	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the public health sector					
Comment:							
1.3	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the private health sector					
Comment:							
1.4	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the private health sector					
Comment:							
2	Practicability	The maturity assessment method of the HCDMMM can be put into action to assess the maturity of a health care entity's data management					
Comment:							

Figure I.4: Validation questionnaire *continued*

3.1	Usability	Managers of health care entities will find it easy to use the HCDMMM for assessing data management maturity					
Comment:							
3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment					
Comment:							
4.1	Strengths	What, in your view, are the key strengths of the HCDMMM?					
Comment:							
4.2	Weaknesses	What, in your view, are the key weaknesses of the HCDMMM?					
Comment:							
4.3	Weaknesses	If the HCDMMM was to fail to achieve its stated aim, what do you think would be the reason for this?					
Comment:							

I.2 SMEs' responses

This section consists out of the different SMEs' responses as they completed the validation questionnaire.

I.2.1 Validation SME 1 responses

Figure I.5: Validation SME 1 answers

Validation questionnaire

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
			5	4	3	2	1
1.1	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the public health sector		x			
Comment:							
1.2	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the public health sector	x				
Comment:							
1.3	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the private health sector			x		
Comment: I would prefer to not comment based on assumption. My experience resides within the public domain.							
1.4	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the private health sector			x		
Comment: I would prefer to not comment based on assumption. My experience resides within the public domain.							
2	Practicability	The maturity assessment method of the HCDMMM can be put into action to assess the maturity of a health care entity's data management	x				
Comment:							

Figure I.6: Validation SME 1 answers *continued*

3.1	Usability	Managers of health care entities will find it easy to use the HCDMMM for assessing data management maturity		x			
Comment:							
3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment		x			
Comment: The robot colours for the overview of the results clearly illustrate where improvement is required. I would consider adding an index with a short description of what the red, orange and green ranges indicate.							
4.1	Strengths	What, in your view, are the key strengths of the HCDMMM?					
Comment: The HCDMMM framework is not limited to electronic data, but also considers other data formats, which will be advantageous in the public domain. The model is easy to understand, the instructions are clear and the dropdown arrows will ensure that only valid counts are captured (data validation). The capability area vs domain component summary provides excellent visibility on which key areas to focus on.							
4.2	Weaknesses	What, in your view, are the key weaknesses of the HCDMMM?					
Comment: The HCDMMM framework requires broad data management knowledge, which will most likely require more that one individual's input. In the public domain (specifically rural areas) computers with Excel is not a given and will have to be provided to use the model.							
4.3	Weaknesses	If the HCDMMM was to fail to achieve its stated aim, what do you think would be the reason for this?					
Comment: The success of the HCDMMM will be determined by the improvements applied to the underperforming areas that were identified. In the public domain, health care facilities and organisations are overwhelmed and understaffed. Suggested improvements can fall through the cracks if not followed up on. The HCDMMM may fail to achieve its aim if the action on outcomes are not clear.							

I.2.2 Validation SME 2 responses

Figure I.7: Validation SME 2 answers

Validation questionnaire

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
			5	4	3	2	1
1.1	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>facility</u> level health care entities in the <u>public health sector</u>			x		
Comment: I think it would be more beneficial to tailor it specifically to a public health system, as the “organisation level” is the national government and the facility level is their facilities. In Zimbabwe at present there is not really an electronic patient record system (they utilise exercise books); the data management system is for capturing country wide indicators which are then used to develop reports and inform decision making. The facilities are completely at the mercy of the organisational decisions and struggle with the capacitation of the HCWs and the existing infrastructure.							
1.2	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>organisational</u> level health care entities in the <u>public health sector</u>			x		
Comment: The organisational level does not necessarily collect any data but determines which data the facilities will collect and then pulls these data sets for their use. The facilities largely do not use the data which they collect.							
1.3	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>facility</u> level health care entities in the <u>private health sector</u>					
Comment: N/A (In my setting and area of expertise, public health in Zimbabwe, I cannot give an informed position- although I do think it would be accurate.)							
1.4	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>organisational</u> level health care entities in the <u>private health sector</u>					
Comment: N/A (In my setting and area of expertise, public health in Zimbabwe, I cannot give an informed position- although I do think it would be accurate.)							
2	Practicability	The maturity <u>assessment method</u> of the HCDMMM can be <u>put into action</u> to <u>assess</u> the maturity of a health care entity's data management	x				

Figure I.8: Validation SME 2 answers *continued*

Comment: Yes, after being tailored to suit the system being assessed. As mentioned there are two reasons for data collection 1) patient data and medical history 2) data points for reporting. Which are captured differently for different purposes and different audiences. Once the system being analysed is established this would be incorporated into the tool and the tool would be a great assessment.

3.1	Usability	Managers of health care entities will find it easy to use the HCDMMM for assessing data management maturity	x				

Comment:
Yes, as mentioned I think it would need to be tailored to suit the specific audience, but once this is done I think it would be a very easy to use tool.

3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment	x				

Comment: Yes very!

4.1	Strengths	What, in your view, are the key strengths of the HCDMMM?

Comment:
The radar/web chart are an excellent visualisation tool, that allows the users to clearly see the outcomes and would be very useful to insert into feedback sessions/reports when writing up the Maturity Model! The excel sheet is well laid out, practical and easy to use.

4.2	Weaknesses	What, in your view, are the key weaknesses of the HCDMMM?

Comment:
I appreciate it is a generic tool, but in my experience I think it would be more useful to cater for a more specific audience- such as private or public as these are quite different entities that function in very different ways. This being said, it could be utilised as ground work to tailor to a specific setting.

4.3	Weaknesses	If the HCDMMM was to fail to achieve its stated aim, what do you think would be the reason for this?

Comment: Perhaps that the tool is to generic, or that the data systems in developing countries (beyond South Africa- like Zimbabwe) are almost too behind for such a tool to be applicable yet. As mentioned we still mainly use paper based systems and are not yet capturing patient data but merely trying to get accurate capture of basic indicators right, with our main challenges being staffing, skill and infrastructure.

I.2.3 Validation SME 3 responses

Figure I.9: Validation SME 3 answers

Validation questionnaire

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
			5	4	3	2	1
1.1	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the public health sector	X				
Comment: Yes, it captures the main domains of data management and the domains are applicable at both facility level and organisational level.							
1.2	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the public health sector	X				
Comment:							
1.3	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the private health sector	X				
Comment:							
1.4	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the private health sector		X			
Comment:							
2	Practicability	The maturity assessment method of the HCDMMM can be put into action to assess the maturity of a health care entity's data management	X				
Comment: Yes, the model can be used to assess the current state of maturity of both health facilities and organisational levels							
3.1	Usability	Managers of health care entities will find it easy to use the HCDMMM for assessing data management maturity		X			
Comment: Yes, they can. It looks quite comprehensive. One drawback is that it might be time consuming to complete							

Figure I.10: Validation SME 3 answers *continued*

3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment	X				

Comment: Radar diagrams are good visual representative tools.

4.1	Strengths	What, in your view, are the key strengths of the HCDMMM?

Comment: Addresses all the dimensions for data management
Facilities will be able to see their current state and hence focus on an improvement strategy/roadmap

Visual display of results does assist executives who do not have time to go through the model. The radar graphs show the low hanging fruits that can be addressed with quick returns and impact on the performance of an organisation

Simple / no complexity in ranking maturity states

Very comprehensive and informative

4.2	Weaknesses	What, in your view, are the key weaknesses of the HCDMMM?

Comment: Might be time consuming and requires effort

For tools to be adopted for use, there should be a clear link between how they will impact performance of organisations and facilities

4.3	Weaknesses	If the HCDMMM was to fail to achieve its stated aim, what do you think would be the reason for this?

Comment: Question: Is it a self-assessment model or it needs a consultant to help them do an assessment.
1. If it is a self-assessment model, users might find it difficult to understand or interpret some of the concepts terms especially if they are not from a technical background.
2. It requires someone with knowledge to take managers through the model
3. The weakness with most maturity models is that they are long and too comprehensive, hence becoming time consuming as well as requiring a lot of effort to use

I.2.4 Validation SME 4 responses

Figure I.11: Validation SME 4 answers

Validation questionnaire

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
			5	4	3	2	1
1.1	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>facility</u> level health care entities in the <u>public health sector</u>	X				
Comment:							
1.2	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>organisational</u> level health care entities in the <u>public health sector</u>	X				
Comment:							
1.3	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>facility</u> level health care entities in the <u>private health sector</u>	X				
Comment:							
1.4	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>organisational</u> level health care entities in the <u>private health sector</u>	X				
Comment:							
2	Practicability	The maturity <u>assessment method</u> of the HCDMMM can be <u>put into action</u> to <u>assess</u> the maturity of a health care entity's data management		X			
Comment: See comments in 5.3							

Figure I.12: Validation SME 4 answers *continued*

3.1	Usability	Managers of health care entities will find it easy to use the HCDMMM for assessing data management maturity		X			

Comment: See comments in 5.3

3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment		X			

Comment: See comments in 5.3

4.1	Strengths	What, in your view, are the key strengths of the HCDMMM?

Comment:
Easy to understand and use;
No need for internet connection;
Can be used on low-performance devices;
Could be used by organizations such as WHO or the Global Fund (together with country facilities) to gather information on how facilities manage their data in developing countries. I suspect they will find trends and patterns across facilities and countries. These trends can be used to assess which capability areas, in general, need improvement and will be useful for such organizations to invest and improve on.

4.2	Weaknesses	What, in your view, are the key weaknesses of the HCDMMM?

Comment:
None

4.3	Weaknesses	If the HCDMMM was to fail to achieve its stated aim, what do you think would be the reason for this?

Comment:
In many developing countries, appointed data managers are not always educated or xxx for the role. In these cases, they might not comprehend the reason for using the model in the 1st place. Similarly, they might also lack the knowledge and skills required to accurately use the model – be it, not understanding the descriptions for scoring the capability areas, or interpreting the results.

It is also possible that some data managers might not be honest when evaluating the capability areas, thus making the results useless.

I.2.5 Validation SME 5 responses

Figure I.13: Validation SME 5 answers

Validation questionnaire

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
			5	4	3	2	1
1.1	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>facility</u> level health care entities in the <u>public health sector</u>					

Comment:
Having spent some time working in the public healthcare sector, I envisage that knowledge on how to complete this assessment may be limited. We also operate in a relatively "technology naive" environment where exposure to electronic data capture mechanisms, data storage etc. has not had to be dealt with on a large scale, particularly by smaller sites.

Not enough experience to rank.

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
1.2	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>organisational</u> level health care entities in the <u>public health sector</u>					

Comment:
Not enough experience to comment / rank.

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
1.3	Applicability	The domain components and capability areas of the HCDMMM are <u>representative</u> of real-world <u>facility</u> level health care entities in the <u>private health sector</u>		x			

Comment:
The applicability of the model will depend on the size of the private healthcare facility or group that it belongs to. If part of a bigger enterprise structure a number of these elements will not be in control of the facility with only the capture processes residing at facility level and the remaining part of the process lying within centralised structures.

Given that all doctors work independently within the private healthcare sector however, they will often have their own systems and structures in place in which case this may then be relevant to individual providers also who would work within a facility.

In line with the above comment, the definition of facility could then be expanded to include the specific facility (e.g. Hospital X) or a practice within a facility. Examples being you could have a hospital that has its own data management process and structure in place but the radiology services, pathology services, different specialists, the emergency centre would all be separate and independently operating entities within the facility. This is one key difference between the public and private healthcare

Figure I.14: Validation SME 5 answers *continued*

operating models that currently exists.

1.4	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the private health sector		x			

Comment:
Referring to the field data capturing or storage for management and decision making:
Often data may not always be collected for the primary purpose of management or decision making. Data collected in other operational processes, either clinical, financial, administrative etc, could then be reported and used by organisations for management and decision making. Bedside decision making of a HCP and operational decision making using trend analysis of clinical or operational metrics is also somewhat different. If care coordination is referred to in the data usage component I would assume both clinical decision making at the bedside, along with quality assurance and improvement trend reporting etc. are being included here.

The data types field I would be interested to know why (2) contains structured data only. I know of some sites in the public sector that have scanned the patient files and although that creates an electronic record, this data is unstructured with limited OCR and doesn't allow for extended analytics.

In data sharing and storage there is a contractual component particularly when third parties are involved that should be addressed from a POPIA and in our case GDPR perspective.
There is also a consent consideration that needs to be in place for any demographic or clinical information processing or storage. This would be applicable across sectors. This may reside under alignment to policies and regulations however if you have an individual completing this who is not completely aware of those regulations this could get missed.

2	Practicability	The maturity assessment method of the HCDMMM can be put into action to assess the maturity of a health care entity's data management			x		

Comment:
Identifying the correct person to do this that has a view across multiple business areas or departments would be important to get an accurate result.
If doing this on an organisation level I would assume you would need to answer a question based on having frame of reference of "in the majority of datasets." in mind.

3.1	Usability	Managers of health care entities will find it easy to use the HCDMMM for assessing data management maturity		x			

Comment:
The flow of the tool is easy to understand.

3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment		x			

Comment:
Having been involved in the development and roll out of reports in the different HC sectors, many people struggle interpreting data and reports. In the reporting component, although it may seem very obvious and trivial to someone with any basic stats or excel knowledge, I would add detail on what a red, orange and green suggests or even what

Figure I.15: Validation SME 5 answers *continued*

you should look out for on a spider graph e.g. This pattern = good. This pattern = concern. I would also remove abbreviations of SS, PA etc and replace them with their actual names and abbreviations in brackets. If you can avoiding a user having to cross reference back to another sheet or information source in surfacing the info it often allows for a better experience and limits confusion or misinterpretation.

4.1	Strengths	What, in your view, are the key <u>strengths</u> of the HCDMMM?

Comment:
It covers key areas pertinent to healthcare data. It's easy to complete. It provides a graphic of the reported outcomes.

4.2	Weaknesses	What, in your view, are the key <u>weaknesses</u> of the HCDMMM?

Comment:
Given the vast number of systems that are used in healthcare operations having a high level view on data management maturity doesn't provide direction on where a specific quality gap exists and would still require further work to get to the specific area needing to be addressed. Practically also business ownership of different systems resides in different departments- e.g. Administration, finance, clinical, HR etc. You would then likely need someone who understands the specific business process impact of that area to assess that domain or alternatively the completion of the maturity model would need to be done by a specific assigned/ trained resource who can assimilate the info.

I haven't had exposure to reams of other maturity models so I'm not sure if the following point would reside in the scope of this model. Regardless of sector or entity type, there is often a large people and workflow/ business process component to attaining data. As the industry moves to electronic health record keeping etc. it may be worthwhile considering whether there is a specific socio-technical view on workflow and processes that could be accounted for in the model.

4.3	Weaknesses	If the HCDMMM was to <u>fail</u> to achieve its stated aim, <u>what</u> do you think would be the reason for this?

Comment:
In order to allow for more granular view this could be done across business units within an organisation – adding a reporting layer of those units mapped against each other would then also be beneficial.

In terms of being used for improvement: My concern would be that people wouldn't know what to do with the report and where to start.
In order to give further guidance on what an output means I would potentially supplement the report output with something showing "what the result means" if the average is 1. 2, 3 etc.

We currently already do in depth privacy and security assessments on each individual system – there may then be some redundancy in my specific work environment although I wouldn't see this as a reason to not use it in totality.

I.2.6 Validation SME 6 responses

Figure I.16: Validation SME 6 answers

Validation questionnaire

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
			5	4	3	2	1
1.1	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the public health sector		x			
Comment: What about alignment with intervention objectives? The purpose of the data gathering process will make a huge difference to what is done; also the use of the data feeds back into how this is gathered. E.g. if managers don't use data to help employees fix the very things that they generate the data for (e.g. stock-outs)– they will stop collecting this.							
1.2	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the public health sector		x			
Comment: Same point as above							
1.3	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the private health sector		x			
Comment: Same point as above							
1.4	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the private health sector		x			
Comment: Same point as above							
2	Practicability	The maturity assessment method of the HCDMMM can be put into action to assess the maturity of a health care entity's data management		x			
Comment: There are at present many maturity models being developed – it seems that we need to assess and take stock of maturity on a wide range of areas; I would however have suggested if I saw this earlier that you should have evaluated the importance and implementation difficulty of the elements in the model – to evaluate a model on this level is somewhat in direct to the specific components.							

Figure I.17: Validation SME 6 answers *continued*

3.1	Usability	Managers of health care entities will find it easy to use the HCDMMM for assessing data management maturity		x			
Comment: This too is user friendly – I think it is well laid out and will be easy to use							
3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment		x			
Comment: Some visualisations have been developed – this will help a lotto interpret the maturity assessment							
4.1	Strengths	What, in your view, are the key strengths of the HCDMMM?					
Comment: Clarity and simplicity							
4.2	Weaknesses	What, in your view, are the key weaknesses of the HCDMMM?					
Comment: Not sure of this has been evaluated on a component level? As a whole it look good – but there could be more insight developed if you asked to be ranked on component level; Some aspects are more difficult to get right than others; The elements are also on different levels – so what do you exactly mean with "Data collection technology"; This is very vague and can be many different technologies for different objectives.							
4.3	Weaknesses	If the HCDMMM was to fail to achieve its stated aim, what do you think would be the reason for this?					
Comment: I think this is best answered by applying this to a case study; I have not applied this to a specific problem – so if time permits / if you have done this already – it may be a good idea to do a case study							

I.2.7 Validation SME 7 responses

Figure I.18: Validation SME 7 answers

Validation questionnaire

Question no	Validation target	Statement or question	Strongly agree	Agree	Unsure	Disagree	Strongly disagree
			5	4	3	2	1
1.1	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the public health sector		x			
Comment:							
1.2	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the public health sector		x			
Comment:							
1.3	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world facility level health care entities in the private health sector		x			
Comment:							
1.4	Applicability	The domain components and capability areas of the HCDMMM are representative of real-world organisational level health care entities in the private health sector		x			
Comment:							
2	Practicability	The maturity assessment method of the HCDMMM can be put into action to assess the maturity of a health care entity's data management			x		
Comment:							

Figure I.19: Validation SME 7 answers *continued*

3.1	Usability	Managers of health care entities will find it easy to use the HCDMMM for assessing data management maturity			x		

Comment:

3.2	Usability	It is easy to interpret the results of the HCDMMM's maturity assessment			x		

Comment:

4.1	Strengths	What, in your view, are the key strengths of the HCDMMM?

Comment:
Governance and oversight

4.2	Weaknesses	What, in your view, are the key weaknesses of the HCDMMM?

Comment:
Health data operates from and originates over 50 differing systems and thus aggregated and scrubbed results need to be considered to ascertain the true median / indicator evident.

Mgt should also only focus on exception data / reports

4.3	Weaknesses	If the HCDMMM was to fail to achieve its stated aim, what do you think would be the reason for this?

Comment:
I think it needs to be more user friendly to complete:

Using more pictures etc. diagrams.

List of references

Aanestad, M., Joliffe, B., Mukherjee, A. and Sahay, S. (2014). Infrastructuring Work: Building a state-wide hospital information infrastructure in India. *Information Systems Research*, vol. 25, no. 4, pp. 834–845.

Adams, B. (2017). The importance of data management in healthcare. Information Builders. [online] Available at: `https://www.informationbuilders.com/blog/beth-adams/importance-data-management-healthcare` [Accessed 10/05/2019].

Alkraiji, A.I., Jackson, T. and Murray, I. (2016). Factors impacting the adoption decision of health data standards in tertiary healthcare organisations in saudi arabia. *Journal of Enterprise Information Management*, vol. 29, no. 5, pp. 650–676.

Allorto, N. and Wise, R. (2015). Development and evaluation of an integrated electronic data management system in a south african metropolitan critical care service. *Southern African Journal of Anaesthesia and Analgesia*, vol. 21, no. 6, pp. 31–35.

Bahri, S., Zoghlami, N., Abed, M. and Tavares, J.M.R. (2018). Big data for healthcare: A survey. *IEEE Access*, vol. 7, pp. 7397–7408.

Becker, J., Knackstedt, R. and Pöppelbuß, J. (2009). Developing maturity models for it management. *Business & Information Systems Engineering*, vol. 1, no. 3, pp. 213–222.

Bergum, B.-i., Nielsen, P. and Sæbø, J.I. (2017). Patchworks of Logistics Management Information Systems : Challenges or Solutions for Developing Countries? In: *International conference on social implications of computers in developing countries*, pp. 47–58.

Bhadani, A.K. and Jothimani, D. (2016). Big data: challenges, opportunities, and realities. In: *Effective Big Data management and opportunities for implementation*, pp. 1–24. IGI Global.

Bhaskaran, S., Suryanarayana, G., Basu, A. and Joseph, R. (2013). Cloud-enabled search for disparate healthcare data: a case study. In: *2013 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*, pp. 1–8. IEEE.

Bijlmakers, L., Mueller, D., Kahveci, R., Chen, Y. and van der Wilt, G.J. (2017). Integrate-Hta: a Low- and Middle-Income Country Perspective. *International Journal of Technology Assessment in Health Care*, vol. 33, no. 5, pp. 599–604.

Bindle, S. and Joost, G. (2010). Modeling access to healthcare Design: Design research in remote communities Information. *Information Design Journal*, vol. 18, no. 3, pp. 262 – 274.

Blondiau, A., Cleven, A. and Rohner, P. (2013). Towards a hospital cooperation maturity model-construction and evaluation of a maturity model and a corresponding tool for the identification of challenges and success factors for cooperation in the hospital sector. *International Journal of Health Information Management Research*, vol. 1, no. 1, pp. 4–14.

Braa, J. and Sahay, S. (2012). Improving quality and use of data through data-use workshops : Zanzibar , United Republic of Tanzania. *Bulletin of the World Health Organization*, vol. 90, no. 1, pp. 379–384.

Burnham, J.F. (2006). Scopus database: a review. *Biomedical Digital Libraries*, vol. 3, no. 1, p. 1.

Carvalho, J.V., Rocha, Á. and Abreu, A. (2016). Maturity models of healthcare information systems and technologies: a literature review. *Journal of medical systems*, vol. 40, no. 6, p. 131.

Carvalho, J.V., Rocha, Á., van de Wetering, R. and Abreu, A. (2019*a*). A maturity model for hospital information systems. *Journal of Business Research*, vol. 94, pp. 388–399.

Carvalho, J.V., Rocha, Á., Vasconcelos, J. and Abreu, A. (2019*b*). A health data analytics maturity model for hospitals information systems. *International Journal of Information Management*, vol. 46, pp. 278–285.

Cavanillas, J.M., Curry, E. and Wahlster, W. (2016). *New horizons for a data-driven economy: a roadmap for usage and exploitation of big data in Europe*. Springer.

Chen, G.K.C. (1975). What Is the Systems Approach? *Interfaces*, vol. 6, no. 1, pp. 32–37.

Chimbari, M.J. (2017). Lessons from implementation of ecohealth projects in southern africa: A principal investigator's perspective. *Acta tropica*, vol. 175, pp. 9–19.

Cline, G.B. and Luiz, J.M. (2013). Information technology systems in public sector health facilities in developing countries : the case of South Africa. *BMC medical informatics and decision making*, vol. 13, no. 1, p. 13.

Conwell, C.L., Enright, R. and Stutzman, M.A. (2000). Capability maturity models support of modeling and simulation verification, validation, and accreditation. In: *2000 Winter Simulation Conference Proceedings (Cat. No. 00CH37165)*, vol. 1, pp. 819–828. IEEE.

Dallemule, L. and Davenport, T.H. (2017). What's your data strategy? *Harvard Business Review*, pp. 114–121.

De Bruin, T., Rosemann, M., Freeze, R. and Kaulkarni, U. (2005). Understanding the main phases of developing a maturity assessment model. In: *Australasian Conference on Information Systems (ACIS):*, pp. 8–19. Australasian Chapter of the Association for Information Systems.

Directorate Informatics (). Informatics Capability Maturity Model (ICMM). [online] Available at: `https://drive.google.com/file/d/0B6dKdxaNjBENeTZnaWpoaE1IZTg/view` [Accessed 21/02/2020].

Drucker, P.F. (1988). The coming of the new organization.

Engelbrecht, B. and Edwards, P. (). Data analysis and design method framework. [onliine] Available at: `https://55a4be53-e17c-4adc-9b58-cfd5dcc843bc.filesusr.com/ugd/265383{_}6b568b93436b402ea897fa88a79e1948.pdf` [Accessed: 21/05/2020].

Evariant (2019). What is Healthcare Data Management and Why is It Important? [online] Available at: `https://www.evariant.com/faq/why-is-healthcare-data-management-important` [Accessed 13/05/2019].

Falagas, M.E., Pitsouni, E.I., Malietzis, G.A. and Pappas, G. (2008). Comparison of pubmed, scopus, web of science, and google scholar: strengths and weaknesses. *The FASEB journal*, vol. 22, no. 2, pp. 338–342.

Ferlie, E.B. and Shortell, S.M. (2001). Improving the quality of health care in the united kingdom and the united states: a framework for change. *The Milbank Quarterly*, vol. 79, no. 2, pp. 281–315.

Fitterer, R. and Rohner, P. (2010). Towards assessing the networkability of health care providers: a maturity model approach. *Information Systems and E-Business Management*, vol. 8, no. 3, pp. 300–333.

Flott, K., Callahan, R., Darzi, A. and Mayer, E. (2016). A patient-centered framework for evaluating digital maturity of health services: a systematic review. *Journal of medical Internet research*, vol. 18, no. 4, p. e75.

Franz, M., de Visser, I., Huppertz, A., Bloomfield, E., Corbyn, D. and Osorio-Cortes, L.E. (2015). Building energy access markets: A value chain analysis of key energy market systems.

Fraser, P., Moultrie, J. and Gregory, M. (2002). The use of maturity models/grids as a tool in assessing product development capability. In: *IEEE international engineering management conference*, vol. 1, pp. 244–249. IEEE.

Fritz, F., Tilahun, B. and Dugas, M. (2015). Success criteria for electronic medical record implementations in low-resource settings : a systematic review. *Journal of the American Medical Informatics Association*, vol. 22, no. 2, pp. 479–488.

Frost & Sullivan (2015). Healthcare Data Breaches: Continual Security Improvement as a Path to Lower Risk. Tech. Rep..

Gainer, A., Roth, M., Strong, P. and Davis, J. (2012). A standards-based open source application to gather health assessment data in developing countries. In: *2012 IEEE Global Humanitarian Technology Conference*, pp. 293–298. IEEE.

Galetto, M. (2016). What is data management? [online] NGDATA. Available at: `https://www.ngdata.com/what-is-data-management/` [Accessed: 10/05/2019].

Gandomi, A. and Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International journal of information management*, vol. 35, no. 2, pp. 137–144.

Ganiga, R., Pai, R.M., MM, M.P. and Sinha, R.K. (2018). Private cloud solution for securing and managing patient data in rural healthcare system. *Procedia Computer Science*, vol. 135, pp. 688–699.

Grindle, M., Kavathekar, J. and Wan, D. (2013). A new era for the healthcare industry - Cloud computing changes the game. [online] Accenture. Available at: `http://www.accenture.com/us-en/Pages/insight-healthcare-industry-cloud-computing.aspx` [Accessed 10/05/2019].

Haskew, J., Gunnar, R., Turner, K., Kimanga, D. and Sirengo, M. (2015). Implementation of a Cloud-Based Electronic Medical Record to Reduce Gaps in the HIV Treatment Continuum in Rural Kenya. *Plos One*, vol. 10, no. 8, p. e0135361.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004). Design science in information systems research. *MIS quarterly*, pp. 75–105.

Huang, L., Bai, X. and Nair, S. (2008). Developing a sse-cmm-based security risk assessment process for patient-centered healthcare systems. In: *Proceedings of the 6th international workshop on Software quality*, pp. 11–16.

Idoga, P.E., Toycan, M., Nadiri, H. and Çelebi, E. (2018). Factors affecting the successful adoption of e-health cloud based health system from healthcare consumers' perspective. *IEEE Access*, vol. 6, pp. 71216–71228.

Impact Advisors (2015). Healthcare Information Security Adoption Model (HISAM). [online] Available at: `https://www.impact-advisors.com/insights-impact/hisam/` [Accessed 20/02/2020].

Irawan, Y.S., Soegijoko, S., Koesoema, A.P., Utama, D.Q., Riyani, A., Isdiningrat, A.A., Isdiningrat, I.S. and Husin, F. (2016). Towards sustainable mhealth applications for maternal and child health: The case of sahabat bundaku-an integrated mobile application for mothers and midwives. pp. 3270–3274.

Jaana, M., Paré, G. and Sicotte, C. (2009). Information technology capacities assessment tool in hospitals: Instrument development and validation. *International journal of technology assessment in health care*, vol. 25, no. 1, pp. 97–106.

Jackson, S., Consulting, B.S. and Eisner, H. (2010). What Is the Systems Approach? *ResearchGate*.

Jobson, G., Murphy, J., Van Huyssteen, M., Myburgh, H., Hurter, T., Grobbelaar, C.J., Struthers, H.E., McIntyre, J.A. and Peters, R.P. (2018). Understanding health worker data use in a south african antiretroviral therapy register. *Tropical Medicine & International Health*, vol. 23, no. 11, pp. 1207–1212.

Johnston, D.S. (2017). Digital maturity: are we ready to use technology in the nhs? *Future healthcare journal*, vol. 4, no. 3, p. 189.

Jones, M., Koziel, C., Larsen, D., Berry, P. and Kubatka-Willms, E. (2017). Progress in the enhanced use of electronic medical records: data from the ontario experience. *JMIR medical informatics*, vol. 5, no. 1, p. e5.

Kaposhi, B.M., Mqoqi, N. and Schopflocher, D. (2014). Evaluation of antiretroviral treatment programme monitoring in Eastern Cape , South Africa. *Health policy and planning*, vol. 30, no. 5, pp. 547–554.

Kaseke, F., Stewart, A., Gwanzura, L., Hakim, J. and Chikwasha, V. (2017). Clinical characteristics and outcomes of patients with stroke admitted to three tertiary hospitals in zimbabwe: A retrospective one-year study. *Malawi Medical Journal*, vol. 29, no. 2, pp. 177–182.

Khan, S.I. and Hoque, A.S.L. (2016*a*). Privacy and security problems of national health data warehouse: a convenient solution for developing countries. In: *2016 International Conference on Networking Systems and Security (NSysS)*, pp. 1–6. IEEE.

Khan, S.I. and Hoque, A.S.M.L. (2016*b*). An analysis of the problems for health data integration in bangladesh. In: *2016 International Conference on Innovations in Science, Engineering and Technology (ICISET)*, pp. 1–4. IEEE.

Khan, S.I. and Hoque, A.S.M.L. (2017). Health data integration with secured record linkage: A practical solution for bangladesh and other developing countries. pp. 156–161.

Kleynhans, E. (2020). *The development of a policy analysis logic model to support public medicine availability initiatives in the context of VAN.* Ph.D. thesis, University of Stellenbosch, South Africa.

Koivu, A., Mavengere, N., Ruohonen, M.J., Hederman, L. and Grimson, J. (2016). Exploring the information and ict skills of health professionals in low-and middle-income countries. In: *International Conference on Stakeholders and Information Technology in Education*, pp. 152–162. Springer.

Labrinidis, A. and Jagadish, H.V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, vol. 5, no. 12, pp. 2032–2033.

Lahrmann, G., Marx, F., Winter, R. and Wortmann, F. (2011). Business intelligence maturity: Development and evaluation of a theoretical model. In: *2011 44th Hawaii International Conference on System Sciences*, pp. 1–10. IEEE.

Letebo, M. and Shiferaw, F. (2016). Adapting hiv patient and program monitoring tools for chronic non-communicable diseases in ethiopia. *Globalization and health*, vol. 12, no. 1, p. 26.

Levy, E. (2018). 5 Essential Components of a Data Strategy. [online] SAS Institute Inc. Available at: `https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/5-essential-components-of-data-strategy-108109.pdf` [Accessed 09/10/2019].

Liu, C.-F., Hwang, H.-G. and Chang, H.-C. (2011). E-healthcare maturity in taiwan. *Telemedicine and e-Health*, vol. 17, no. 7, pp. 569–573.

Ludwick, D.A. and Doucette, J. (2009). Adopting electronic medical records in primary care: lessons learned from health information systems implementation experience in seven countries. *International journal of medical informatics*, vol. 78, no. 1, pp. 22–31.

Mapowo, N. (2019). Development of a pharmaceutical product categorization framework taxonomy for the south african public healthcare pharmaceutical supply chain-van enabled.

Masana, N. and Muriithi, G.M. (2017). Investigating toe factors affecting the adoption of a cloud-based emr system in the free-state, south africa. In: *International Conference on Emerging Technologies for Developing Countries*, pp. 233–238. Springer.

Mate, K.S., Bennett, B., Mphatswe, W., Barker, P. and Rollins, N. (2009). Challenges for Routine Health System Data Management in a Large Public Programme to Prevent Mother-to-Child HIV Transmission in South Africa. *Plos One*, vol. 4, no. 5, p. e5483.

Medhanyie, A.A., Spigt, M., Yebyo, H., Little, A., Tadesse, K., Dinant, G.-J. and Blanco, R. (2017). Quality of routine health data collected by health workers using smartphone at primary health care in ethiopia. *International journal of medical informatics*, vol. 101, pp. 9–14.

Medhi, I., Jain, M., Tewari, A., Bhavsar, M., Matheke-Fischer, M. and Cutrell, E. (2012). Combating rural child malnutrition through inexpensive mobile phones. In: *Proceedings of the 7th Nordic conference on human-computer interaction: making sense through design*, pp. 635–644.

Mengiste, S.A. (2013). Globalization and icts: Potentials and challenges to the public health sector of of developing countries. *Association for Infomration Systems*.

Mettler, T. (2010*a*). *Supply management in the hospital: construction and evaluation of a configurable maturity level model for targeted design*. Ph.D. thesis, University of St. Gallen, Switzerland.

Mettler, T. (2010*b*). Thinking in terms of design decisions when developing maturity models. *International Journal of Strategic Decision Sciences (IJSDS)*, vol. 1, no. 4, pp. 76–87.

Mettler, T. and Blondiau, A. (2012). Hcmm-a maturity model for measuring and assessing the quality of cooperation between and within hospitals. In: *2012 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 1–6. IEEE.

Mgozi, T. and Weeks, R. (2011). Cloud computing impact on heathcare in South Africa. In: *IAMOT 2015 Conference Proceedings*, pp. 127–143.

Miller, H.G. and Mork, P. (2013). From data to decisions: a value chain for big data. *IT Professional*, , no. 1, pp. 57–59.

Monda, J., Keipeer, J. and Were, M.C. (2012). Data integrity module for data quality assurance within an e-health system in sub-saharan africa. *Telemedicine and e-Health*, vol. 18, no. 1, pp. 5–10.

Mukhopadhyay, A., Xavier, B., Sreekumar, S. and Suraj, M. (2018). Real-time ecg monitoring over multi-tiered telemedicine environment using firebase. In: *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 631–637. IEEE.

Mvelase, P., Dlamini, Z., Dludla, A. and Sithole, H. (2015). Integration of smart wearable mobile devices and cloud computing in south african healthcare. In: *eChallenges e-2015 Conference*, pp. 1–10. IEEE.

Nambiar, R., Bhardwaj, R., Sethi, A. and Vargheese, R. (2013). A look at challenges and opportunities of big data analytics in healthcare. In: *2013 IEEE international conference on Big Data*, pp. 17–22. IEEE.

Naudet, Y. and Chen, D. (2012). ehealth interoperability evaluation using a maturity model. *Quality of Life Through Quality of Information: Proceedings of MIE2012*, vol. 180, p. 333.

Nicol, E., Bradshaw, D., Phillips, T. and Dudley, L. (2013). Human factors affecting the quality of routinely collected data in south africa. In: *Medinfo*, pp. 788–792.

Nyumbeka, D. and Wesson, J. (2014). Using mobile computing to support malnutrition management in south africa. In: *Proceedings of the Southern African Institute for Computer Scientist and Information Technologists Annual Conference 2014 on SAICSIT 2014 Empowered by Technology*, pp. 352–360.

O'Neal, K. (2012). What is the difference between Information Management and Data Management? [online] Available at: `http://www.b-eye-network.com/blogs/oneal/archives/2012/02/what{_}is{_}the{_}dif.php` [Accessed 13/05/2019].

Pahl, C., Zare, M., Nilashi, M., de Faria Borges, M.A., Weingaertner, D., Detschew, V., Supriyanto, E. and Ibrahim, O. (2015). Role of openehr as an open source solution for the regional modelling of patient data in obstetrics. *Journal of biomedical informatics*, vol. 55, pp. 174–187.

Palmer, R.C. and Simms-cendan, J.S. (2012). Implementing an electronic health record as an objective measure of care provider accountability for a resource-poor rural area in the Dominican Republic. In: *IET Conference Publications*, pp. 1–6.

Paré, G. and Sicotte, C. (2001). Information technology sophistication in health care: an instrument validation study among canadian hospitals. *International journal of medical informatics*, vol. 63, no. 3, pp. 205–223.

Patra, M.R., Das, R.K. and Padhy, R.P. (2012). Crhis: cloud based rural healthcare information system. pp. 402–405.

Paul, S., Das Bhattacharya, S., Sudar, A., Patra, D., Majumdar, A.K., Mukhopadhyay, J. and Majumdar, B. (2009). A web-based electronic health care system for the treatment of pediatric HIV. *2009 11th IEEE International Conference on e-Health Networking, Applications and Services, Healthcom 2009*, pp. 175–180.

Paulk, M.C., Curtis, B., Chrissis, M.B. and Weber, C.V. (1993). Capability maturity model, version 1.1. *IEEE software*, vol. 10, no. 4, pp. 18–27.

Pearce, C.M., De Lusignan, S., Phillips, C., Hall, S. and Travaglia, J. (2013). The computerized medical record as a tool for clinical governance in australian primary care. *Interactive journal of medical research*, vol. 2, no. 2, p. e26.

Pickering, C. and Byrne, J. (2014). The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. *Higher Education Research & Development*, vol. 33, no. 3, pp. 534–548.

Pöppelbuß, J. and Röglinger, M. (2011). What makes a useful maturity model? a framework of general design principles for maturity models and its demonstration in business process management. In: *Ecis*, p. 28.

Raghupathi, W. and Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health information science and systems*, vol. 2, no. 1, p. 3.

Reid, P. and Grossman, J. (2005). A framework for a systems approach to health care delivery. *Building a better delivery system: A new engineering/health care partnership. National Academies Press, Washington, DC, États-Unis.*

Reid, P.P., Compton, W.D., Grossman, J.H., Fanjiang, G. *et al.* (2005). Information and communications systems: The backbone of the health care delivery system. In: *Building a Better Delivery System: A New Engineering/Health Care Partnership*. National Academies Press (US).

Rimmer, C., Hagens, S., Baldwin, A. and Anderson, C.J. (2014). Measuring maturity of use for electronic medical records (emrs) in british columbia: the physician information technology office (pito). *Healthc Q*, vol. 17, no. 04, pp. 75–80.

Röglinger, M., Pöppelbuß, J. and Becker, J. (2012). Maturity models in business process management. *Business process management journal*.

Rosemann, M. (2003). Application reference models and building blocks for management and control. In: *Handbook on Enterprise Architecture*, pp. 595–615. Springer.

Sampath, P., Tamilselvi, S., Saravana Kumar, N., Lavanya, S. and Eswari, T. (2017). Diabetic data analysis in healthcare using Hadoop architecture over big data. *International Journal of Biomedical Engineering and Technology*, vol. 23, no. 2/3/4, p. 137.

Sanders, D., Burton, D.A. and Protti, D. (2013). The healthcare analytics adoption model: A framework and roadmap. *Health Catalyst*, vol. 30.

Schurer, M.J. (2020). *Towards the interoperability of spontaneous reporting systems in pharmacovigilance: A maturity model approach with a sociotechnical system focus*. Ph.D. thesis, University of Stellenbosch, South Africa.

Sharifi, M., Jahanbakhsh, M., Tavakoli, N. and Mokhtari, H. (2013). E-Health Implementation Challenges in Iranian Medical Centers: A Qualitative Study in Iran. *Telemedicine and e-Health*, vol. 19, no. 2, pp. 122–128.

Sharma, B. (2008). Electronic healthcare maturity model (ehmm). *Chennai, India: Quintegra*.

South-African Department of Health (2015). mHealth Strategy 2015-2019. pp. 1–20.

Tarhan, A.K., Garousi, V., Turetken, O., Söylemez, M. and Garossi, S. (2020). Maturity assessment and maturity models in health care: A multivocal literature review. *Digital Health*, vol. 6.

Teo, T.S. and King, W.R. (1997). Integration between business planning and information systems planning: an evolutionary-contingency perspective. *Journal of management information systems*, vol. 14, no. 1, pp. 185–214.

Turan, A.H. and Palvia, P.C. (2014). Critical information technology issues in turkish healthcare. *Information & Management*, vol. 51, no. 1, pp. 57–68.

Van Aken, J.E. and Berends, H. (2007). *Problem solving in organizations*. Cambridge University Press.

Van Dyk, L. (2013). *The development of a telemedicine service maturity model*. Ph.D. thesis, Stellenbosch University, Stellenbosch.

Vayena, E., Salathé, M., Madoff, L.C. and Brownstein, J.S. (2015). Ethical challenges of big data in public health. *PLoS Comput Biol*, vol. 11, no. 2, p. e1003904.

Venable, J. and Baskerville, R. (2012). Eating our own cooking: Toward a more rigorous design science of research methods. *Electronic Journal of Business Research Methods*, vol. 10, no. 2.

Venesco (2015). Developing and testing a data management model and maturity scale tailored to improving patient matching accuracy. [online] Available at: `https://www.healthit.gov/sites/default/files/ptmatchwhitepaper.pdf` [Accessed 21/05/2019].

Vidal Carvalho, J., Rocha, Á. and Abreu, A. (2019). Maturity of hospital information systems: Most important influencing factors. *Health informatics journal*, vol. 25, no. 3, pp. 617–631.

Vivanco, J.S. and Oropeza, M.A. (2017). Challenges for using it in mexico's health care industry (aguascalientes méxico case). *Systemics, Cybernetics and Informatics*, vol. 15, no. 7, pp. 57–61.

Vlădescu, C., Galan, A., Olsavszky, V. and Scîntee, S.G. (2009). Romanian health system strategic directions for the next decade. *Italian Journal of Public Health*, vol. 6, no. 1.

vom Brocke, J. (2007). Design principles for reference modeling: reusing information models by means of aggregation, specialisation, instantiation, and analogy. In: *Reference modeling for business systems analysis*, pp. 47–76. IGI Global.

Vreeland, A., Persons, K.R., Primo, H.R., Bishop, M., Garriott, K.M., Doyle, M.K., Silver, E., Brown, D.M. and Bashall, C. (2016). Considerations for exchanging and sharing medical images for improved collaboration and patient care: Himss-siim collaborative white paper. *Journal of digital imaging*, vol. 29, no. 5, pp. 547–558.

Wang, W., Elliot, S., Williams, M.-A. and Wang, B. (2010). An IS contribution to the UN Millennium Development Goals: Next Generation Vaccination Management in the Developing World. In: *ACIS 2010 Proceedings - 21st Australasian Conference on Information Systems*.

Wang, Y., Kung, L. and Byrd, T.A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, vol. 126, pp. 3–13.

Warkulwiz, G., Paul, L. and Mehta, K. (2014). Efficacy of knowledge transfer and exchange between CHWs and community members in low resource settings. *Proceedings of the 4th IEEE Global Humanitarian Technology Conference, GHTC 2014*, pp. 36–40.

Whalen, C., Donnell, D. and Tartakovsky, M. (2014). Supporting research sites in resource-limited settings: Challenges in implementing IT infrastructure. *Journal of acquired immune deficiency syndromes (1999)*, vol. 65, no. 1, p. 544.

Winter, R. (2003). Modelle, techniken und werkzeuge im business engineering. In: *Business Engineering*, pp. 87–118. Springer.

World Health Organization (2008). Framework and standards for country health information systems.

World Health Organization (2014). Health facility & community data toolkit.

World Health Organization (2018). 2018 global reference list of 100 core health indicators (plus health-related sdgs).

Yang, J.J., Li, J., Mulder, J., Wang, Y., Chen, S., Wu, H., Wang, Q. and Pan, H. (2015). Emerging information technologies for enhanced healthcare. *Computers in industry*, vol. 69, pp. 3–11.

Yaqoob, T., Mir, F., Abbas, H., Shahid, W.B., Shafqat, N. and Amjad, M.F. (2017). Feasibility analysis for deploying national healthcare information system (NHIS) for Pakistan. In: *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–6. IEEE.

Zaino, J. (2017). A Chief Data Officer Challenge: Build Your Data Strategy. [online] Available at: `https://www.dataversity.net/chief-data-officer-challenge-build-data-strategy/` [Accessed 24/06/2020].