Care Mining:

An application of Process Mining in Value-Based Healthcare

Strengthening the concept Value Based Healthcare with Process Mining in practice

1st Supervisor:
Marjan Askari
(Utrecht University)

2nd Supervisor:
Jan Martijn van der Werf
(Utrecht University)

Author:
Runa Sewradj
Master of Business Informatics
Student number: 4148428
E-mail: r.s.sewradj@students.uu.nl

Faculty of Science

Department of Information and Computing Sciences

Princetonplein 5, De Uithof

3584 CC Utrecht

Abstract

The Health Care sector is under considerable pressure concerning its financial aspects. Restricted budgets are available for the performance of care, meanwhile the costs of care are rising. Therefore, research in how and where costs in hospitals could be reduced, is required. This would also stimulate the increase of value in care.

Several studies have been performed on how care could be increased. Most of the studies only focused on the financial aspects and quality management facets. These studies were lacking a stepwise, efficient approach to study care processes to increase value.

This thesis is an analysis of the care processes, with the following research question: "*How can Process Mining strengthen Value-Based Health Care?"* The costs in this thesis refer to performance times of care activities.

A well-known concept in health care provides this research with a base and structure to investigate care processes: Value-Based Health Care (VBHC). VBHC is about maximizing the value of care for patients and to reduce care costs. To put VBHC in practice, the method 'Time-Driven Activity-Base Costing' (TD-ABC) was used, to support increasing care value by creating process models of care processes. The TD-ABC method provides seven steps, where one of the steps describes the manual creation of process models. Due to budget issues of hospitals, this part of the method had to be performed more efficiently. Therefore, Process Mining was introduced, focused on extracting knowledge from information systems to analyze performed processes.

In literature, the method 'Process Mining Process' was mentioned, describing the performance of Process Mining in practice, which involves six steps.

The TD-ABC and Process Mining Process steps were combined to create a more effective and efficient method for analyzing care processes: The Care Mining Method. This method consists of five steps: 1) Select a medical condition, 2) Define Care Delivery Value Chain, 3) Develop Process maps ', 4) Perform analyzes, 5) Discussion and Validation.

With literature, the performance of the Care Mining method could not be tested. Therefore, this method was put into practice with two hospital cases. Unfortunately, more cases could not be acquired due to a lack in awareness of the importance of analyzing care processes, research time restrictions or budget limitations. Nevertheless, the results of these cases show that Process Mining could strengthen VBHC due to a decrease of the performance time of analyzing care processes (1), more visual and distinct process models (2) and more accurate and effective maintenance of the process models.

Further research could focus on investigating the resources of medical specialists, together with the throughput times of patients-related data. Financial aspects can thereafter be linked to the processes and total financial costs of throughput times can be determined.

Keywords: Value-Based Health Care, Process Mining, Time-Driven Activity-Based Costing, Health Care Sector

TABLE OF CONTENTS

1 Introduction 8
1.1 Current situation 8
1.2 Problem definition 9
1.3 Problem statement and objective 12
1.4 Research Questions 12
1.5 Thesis overview 13
2 Research Method 16
2.1 Research Approach 16
2.2 Validity 20
2.3 Contributions 21
2.4 Research planning 21
3 Value-Based Health Care 24
3.1 Approaching Value in Health Care 25
3.2 Measurement of value 25
3.3 VBHC method: Time-Driven Activity-Based Costing 26
4 Process Mining 37
4.1 Process Mining technique 37
4.2 Process Mining tools 39
4.3 Process Mining Method 39
4.4 Process Mining Cases 41
5 Care Mining method 43
5.1 Care Mining method Steps 43
5.2 General evaluation of the method 54
6 Application of the method: CASE STUDY 58
6.1 Case: Breast Cancer 58
6.2 Case: Colorectal 70
7 Validation 79
7.1 Validation by the modeler 79
7.2 Validation by an expert 82
8 Discussion 83
8.1 Main findings 83
8.2 Challenges and Limitations 87
8.3 Further work 87
9 Summary 89
10 References 91
11 Glossary 96

12 Appendices 97
12.1 Compute the total costs over each patient's cycle of care 97
12.2 Theoretical model of Case Mammacarcinoom 98

ACKNOWLEDGEMENT

First, I would like to thank my thesis advisor dr. Marjan Askari of the Department of Information and Computing Sciences at University Utrecht. The office door of dr. Askari was always open whenever I ran into a trouble spot or had a question related to my research or writing. She consistently allowed this paper to be my own work, but guided me towards the right direction whenever she thought it was needed.

Also, I would like to thank dr. ir. Jan Martijn van der Werf of the Department of Information and Computing Sciences at University Utrecht as the second reader of this thesis. I am very grateful for his valuable comments.

Further, I am thankful for to the involvement of Zuiver ICT and especially the owner, Rob van de Coevering, supporting me in acquiring hospital cases. Without that support, it would have been difficult for me to meet medical specialists and inquire cases.

I would also like to express my appreciation to several organizations which gave me the opportunity to gain knowledge about my research topics. Especially I would like to thank the hospitals where I could participate in different meetings and discussions, and where I could perform my method in some cases. Without their passionate participation and input I could not have done the validation of the method.

Finally, I must express my very profound gratitude to my parents and to my sister for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Author

Runa Sewradj

1 INTRODUCTION

This thesis presents a research about costs of care cycles directed to patients. The costs refer to the performance time of each activity in the care cycle. The first chapter elaborates on the current situation, problem definition, problem statement and objectives.

1.1 CURRENT SITUATION

Over the last ten years the productivity of care in hospitals has increased, but lack of insights on value of the delivered care still exists (van den Berg et al., 2014). Herewith, an increase of 58,8% in hospitals admissions shows between 2002 and 2012: from 1620 to 2573 admissions per 1000 inhabitants (CBS, 2014c). From a total of €94.2 billion in 2013, €24.8 billion of the healthcare budget was spent on hospitals and medical specialists (CBS, 2014a). Between 2013 and 2014 the Dutch health care system dropped from a 25th to a 40th place compared to the rest of the world(Bloomberg, 2013). This could imply that a decline in efficiency had occurred. Kaplan and Porter (2011) stated that "there is an almost complete lack of understanding of how much it costs to deliver patient care" and "how those costs compare with the outcomes achieved."

Health care costs in The Netherlands are rising relatively to the inhabitants general income (Bloomberg, 2013). According to Bloomberg (2013) The Netherlands has the second highest rise in health care costs in the world. Between 2001 and 2011 an annualized increase in health care expenditure per capita was recorded at 11.21% (Bloomberg, 2013): per year an average increase of 5.5% (van den Berg et al., 2014). Focusing on the costs in 2014, it is stated that 14,3% of the GDP is spent on health care costs (Centraal Bureau Statistiek [CBS], 2014). In the United States health care costs increase even faster in comparison to The Netherlands and the rest of the world. These costs rose with 5.47% against an annualized increase in income per capita of 2.97%, which is lower than the income growth of The Netherlands. In a report of the National Institute for Health and the Environment it is stated that The Netherlands has the highest health care expenditure in percentage of the GDP in Europe (van den Berg et al., 2014). This is caused by the volume growth, particularly in the mental health care sector. Internationally the expenditures of The Netherlands are on average for curative care, while for long-term care the expenditures are relatively high.

Budgets in the health care sector are increasingly limited (Dol, Raap, & Lageman, 2014; van Deen, Esrailian, & Hommes, 2015). In order to deal with this issue hospitals could set up programs to save in their finances, but this is not feasible for the long-term, caused by the fact that the healthcare sector has a large and diverse number of stakeholders: 1) patients, 2) employers, 3) physicians and other care providers, 4) health plans, 5) suppliers of drugs and medical devices, 6) and governments (Dol et al., 2014; Porter, 2007). Each stakeholder has a different perspective about the performances of processes and the amount of money that should be spent on explicit actions. Primarily focusing on the finances could help to reduce the costs of one stakeholder, but will not automatically lead to a reduction for another stakeholder (Porter & Teisberg, 2006). Therefore, it is difficult to reorganize or change certain procedures in the healthcare sector. Nevertheless, one could create more awareness about the current situation, in which all stakeholders have the same goals or a common purpose. In other words, the various stakeholders need to have a measurable and common view of how hospitals and other care institutions should be managed (Kaplan & Porter, 2011). This research will focus on the view of physicians and care givers (from hospitals).

1.2 PROBLEM DEFINITION

Many studies have been conducted to improve the value of care and its processes (van Deen, Esrailian, & Hommes, 2015; Dol et al., 2014). An example is the DOT-system (Diagnosis Treatment Combination towards transparency) which tries to make costs more transparent (Koomans, 2014). This system uses predefined prices of a care demand, which is based on the price of the entire care product. An entire care product is for example the price of a broken arm, which is determined from X-rays to plaster room. Other studies are focused on quality management, which was introduced in the 1980s (Kunst, Lemmink, & Prins, 1996). Quality management became increasingly important due to influences of government and customers, and initiatives of hospital management (Kunst & Lemmink, 2000). Furthermore, quality management was initiated as a result of decreasing budgets and increasing professional standards of management. Resulting, the primary health care processes were re-evaluated (Kunst et al., 1996). The government has since set up initiatives to promote and improve quality management (The Dutch Council for Health and Care, 2013), such as LEAN projects[1] and Business Process Redesign [3], Evidence Based Medicine [4], and Lean [5], among others. According to Van den Berg et al. (2014) transparency of quality is the main topic to concentrate on, when analyzing health care expenditures. Transparency of quality was a shortcoming in 2010 and had a number of improvement points (van den Berg et al., 2014). Through a significant number of projects, considerable attention is paid to the transparency of health care quality in hospitals. Due to all the new insights and projects, which are initiated and set up around transparency, final conclusions still needs to be defined (van den Berg et al., 2014).

Health care issues can be categorized into the following categories (Mcclellan & Mcclellan, 2011; Porter, Pabo, & Lee, 2013): the price of drugs and new technology; the rising uninsured people; the health plans' unnecessary administrative costs; the provider incentives for overtreatment; the lack of consumer responsibility for cost; the slow penetration of information technology. According to Porter (2005), the important issues where the health care sector is coping with are the following:

1. Costs and access to health insurance.
2. Standards for coverage, or the types of care that should be covered by insurance versus being the responsibility of the individual.
3. The structure of health care delivery.

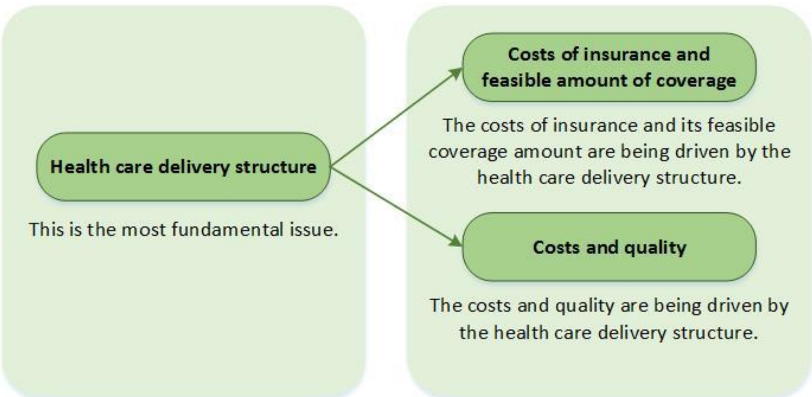


Figure 1: Health care delivery structure as the fundamental issue (Porter, 2005)

[1] http://www.leanforms.com/relaties-aan-het-woord/lean-werkprocessen

Health care delivery structure is the controller of the costs and quality which is being used through the entire health care system (see Figure 1). Controlling health care by using a structure will eventually lead to managing the costs of insurance and the amount of feasible coverage (Ginter, Duncan, & Swayne, 2013).

The aforementioned studies are based on average costs and not on actual costs per individual patient. This is an issue because the costs per individual patient vary considerably (Koomans, 2014; Kaplan et al., 2014). Another issue with these studies is that the caregiver is being paid for performed care instead of on treatment outcomes and increased quality of life. When these issues are taken into account it will help in noticing on which basis decisions are taken and if value is being created for the patient or not (Koomans, 2014). At the end, the focus of hospitals relies on less rework or better cost indication, for which the current processes need to be improved (McLaughlin, Setlur, Kaplan, & Mahajan, 2014).

Currently there is no strong structure in the health care delivery system. This is caused by the fact that there is a lack in competition (Porter, 2010). Competition aims at highlighting and improving quality and costs (Porter, 2010). Innovations and new technologies are distributed in a more accelerated manner. Resulting, the prominent competitors will continue to grow and the weaker competitors are forced to restructure or go out of business. The advantages are that prices on certain treatments are not fixed, but focused on the delivered quality of care, increased value improvement, and the consumers will have more diversity in choosing a health care provider which can meet their expectations. In other sectors, such as mobile communications, consumer banking, computers, etc., the competitors focused-thinking has already been implemented and used successfully (Porter, 2005).

Competition in health care had not have achieved successes as much as it should be, due to unfortunate strategic, organizational, and regulatory choices made by participants in the health care system. The problems in this sector can be allocated to all these participants. Porter (2005) stated the following in accordance to the aforementioned issues: *"[that] the only way to truly reform health care is to reform the nature of competition itself"* (Porter, 2005). Reforming the nature of competition is feasible by focusing on creating competition on a value-based level, competition on a specific medical condition, competition which is around patients, and competition about health results.

The main target in health care is on controlling costs by minimizing it in a short term and preventing it to increase. Most attempts till now have failed to do so. In order to reform health care, its competitive position should be adjusted in accordance with value for patients. For improving value, all the participants need to create a competitive position on a value-based level. Most of the strategies, organizational structures, and practices are poorly aligned with value for patients (Porter, 2009).

Because of the aforementioned arising problems in the health care sector in the US, initiatives started to create theories to achieve these insights. One of these theories is called Value-Based Health Care (VBHC). To create transparency in the processes of hospitals Process Mining is used to extract data from the information systems, which will be explained in the next section.

1.2.1 VALUE-BASED HEALTHCARE

Value-Based Healthcare (VBHC) is a theory developed by Michael Porter (2007) and is focused on the value of care. Value of care can be divided into quality aspects. Quality aspects includes access to care, safety of a treatment, respect as a person, and health outcomes (Porter, 2010; Mcclellan & Mcclellan, 2015). Health outcomes are the measurable factors of value of care when focusing on care processes. The latter could imply whether costs for patients can be reduced and on which parts in the care processes. In this study, we will use the health outcomes to determine whether value for a particular medical condition is being created or not. Value of a patient is measured by dividing health outcomes by its made costs (Porter, 2010).

With a value-focused mind quality and costs are made into key factors. VBHC makes sure that fees (paid by health insurance) are not being paid for the volume of care (that a care-provider gives), but that fees are being paid for the delivered value of care (Hagen, 2015). At the end, value will be created between the patient and their family and the medical specialist and their team (Porter, 2010). This is a new way of approaching the care industry.

1.2.2 PROCESS MINING

Process Mining is a data-driven approach to convert data easily into process information. The goal of Process Mining is to obtain information about the performed processes from an event log (van der Aalst, 2011). The need to use Process Mining can be allocated to the fact that more events or patient-data are being recorded with more detailed information. Also, the need to improve and support the performed processes is a fact (van der Aalst et al., 2012). Hospitals are switching rapidly between their turnover times of activities which will be supported with Process Mining. Herewith, Process Mining makes sure that easy and quick visual overviews are created of current performed processes.

Data used by Process Mining is derived from information systems, such as EHR (Electronic Health Record) systems. By taking the actual data from the information systems of the hospitals, the resulting process-model will be based on real performed care processes. In response to this model, hospitals will get a clear view of their process flow and cost distributions (in time). The latter could be per patient, per patient-group, or per department. At the end, the focus of hospitals relies on less rework or better cost indication, for which the current processes need to be improved (McLaughlin et al., 2014).

Furthermore, the case studies where VBHC together with Process Mining are used in practice are limited. This can be seen by the limited number of hospitals who are in a transition-process of making their activities transparent. The gap between the theory and the practice can endanger the use, development, innovation and adaptation of these two theories. When searching for scientific papers this gap can also be seen. Mans et al. (2008) already stated that “it is not known what happens in a healthcare process for a group of patients with the same diagnosis” (Mans, Schonenberg, Song, & van der Aalst, 2008).

1.3 PROBLEM STATEMENT AND OBJECTIVE

Based on the problem definition explained in the aforementioned section the following problem statement is formulated:

There is insufficient understanding in how value can be created for the patient and what a patient cost in a care cycle for hospitals.

Based on this problem statement, the following objective is formulated:

Provide a method about the effectiveness of combining the theories Value-Based Health Care and Process Mining.

The fundamental question is how to enhance value of health care systems. There are many aspects that a hospital could focus on with regard to increasing value, such as patient satisfaction, enforcing safety regulations, and cost reduction (Boer & Croon, 2008; KPMG, 2013). The main focus should be on creating value in order to review which activities are relevant and of interest for the patient (e.g. a considerably longer life, degree and duration of recovery, better quality of life and less complications). In order to analyze value it is crucial to recognize where value is being created in health care, which is in all care treatment activities that serve a patient's needs (Porter, 2010). Value can be determined by a patient's specific medical condition (Porter, 2010; van Deen, Esrailian, & Hommes, 2015; Petrova, Dale, & Fulford, 2006) (e.g. diabetes, obese, heart diseases) in a full care cycle, namely from monitoring and prevention, treatment and the aftercare. Kaplan & Porter (2011) define a medical condition as: "an interrelated set of patient circumstances that are best addressed in a coordinated way and should be broadly defined to include common complications and comorbidities." A medical condition contains diseases, illnesses, injuries, and natural circumstances (such as pregnancy) (Porter, 2009). This does not include individual procedures, service, office visit, or certain tests. Also, to treat a medical condition care is measured with used capacity. Often hospitals focus on cost reduction efforts like the reduction of expensive treatments, and at the same time denying value that deliver effective patient outcomes (Porter, 2010). In preference, hospitals should invest in activities that relate to primary treatment, such as preventative care and monitoring, and create value in this way. Furthermore, focusing on medical conditions will result in approaching and treating patients more efficiently and effectively, which will lead to an increase in care value. Ultimately, this could, in a way, lead to a decrease in cumulative costs in the whole healthcare cycle.

1.4 RESEARCH QUESTIONS

To structure this research a main research question (MRQ) and several sub-questions (SQ) are formulated. Answering these questions will provide input for physicians and care givers (hospitals), where the topics VBHC together with Process Mining has only gained little attention so far. This is based on the objective formulated in the previous section: "[To] provide a method about the effectiveness of combining the theories Value-Based Health Care and Process Mining". As stated, VBHC includes several quality aspects which can be used to increase value. However, an emphasis on health outcomes is made, since it contains measurable factors. The before mentioned leads to the following questions, starting with the main research question:

The main research question is stated as follows:

MRQ: *"How can Process Mining strengthen Value-Based Health Care?"*

To answer the main research question, the following sub-questions are defined:

SQ1: What is Value-Based Health Care and how are costs related to it?
Sub Research Question 1 (SQ1) focuses on the theoretical findings of the theory Value-Based Health Care and discusses if this theory can be used to enhance care with care costs. With findings, we mean the elements and aspects where the theory of Value-Based Health Care is based upon. Papers with implemented Value-Based Health Care cases are used to analyze cost aspects. By creating more understanding about how costs are related to Value-Based Health Care the relevance of this research increases.

SQ2: What is known about Process Mining in Health Care?
Process Mining is the second theory which will be researched. The theoretical findings in the health care sector will be discussed to create awareness about the theory and its applicability in the care sector. This question needs to clear the relevance of this theory in this research.

SQ3: How can Process Mining support VBHC for identifying and improving costs?
This sub question explains on theoretical level which aspects or elements of Process Mining can be used to link which aspects or elements of Value-Based Health Care. The aspects or elements of Process Mining and Value-Based Health Care need to be cost related in order to create a link for identifying and improving costs. Afterwards, the theory gained from SQ2 and SQ 3 will be used to create a method, which explains how and if Process Mining and Value-Based Health Care can be used together.

SQ4: How can the support of Process Mining within VBHC be put in practice?
After a theoretical bases and method has been build, the created method will be used in real cases to test whether and how it can be used by health care managers. This question will answer the practicability and usability of the method.

Sub-question 1 up to 3 will be answered by the related literature of this thesis and sub-question 4 will be answered by the input of the case studies that are involved in this thesis. One of the most important aspects in this respect is to find out, whether hospitals are able to implement the provided method of this research. The first question substantiates the existing aspects of VBHC, in order to understand the structure of it. For this research, not all the aspects of VBHC will be used. The main focus will be on costs. Therefore, the first sub-question will also support in creating a method. Sub-question 2 will discuss the notion of Process Mining with its aspects and in what way it can be used to analyze data from applications. The third sub-question creates a basis to answer the main research question and will provide an argumentation what the usefulness of this research is. It will describe which aspects of Process Mining can be used to build a bridge with which aspects of VBHC. An overview of the different chapters and the research questions they answer, can be seen in Figure 2.

1.5 THESIS OVERVIEW

The subsequent chapter describes the research approach of this thesis. Next, the related literature will be discussed and is elaborated on in the Background of Chapter 3. In Chapter 4 the composed method will be proposed. Chapter 5 presents the results of the conducted case studies. The primary results will be summarized and validated in Chapter 6. Finally, Chapter 7 will elaborate on the limitations to this research and the last chapter of this thesis will be discussing the results of the performed case studies in Chapter 8.

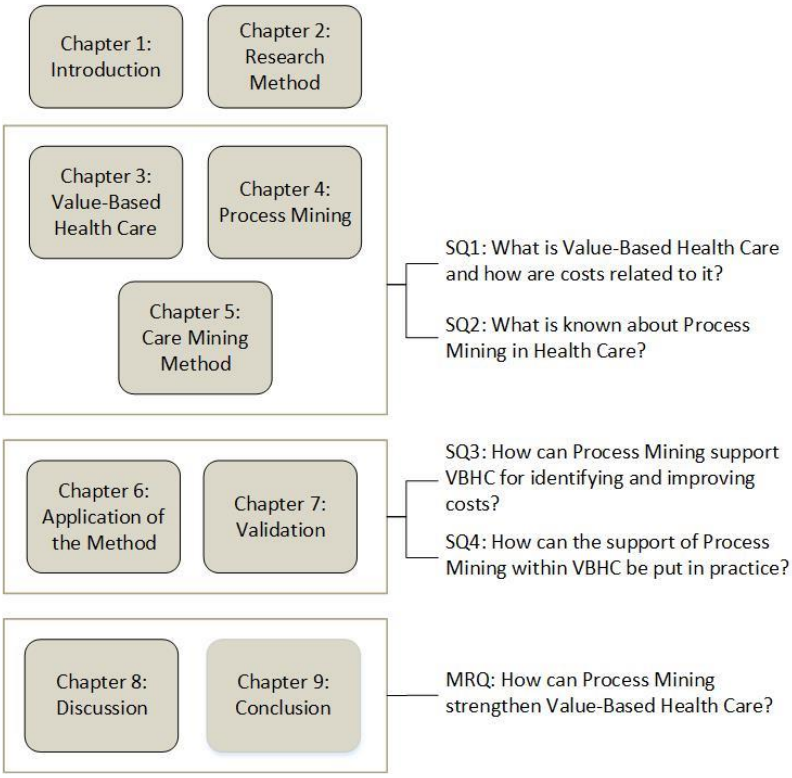


Figure 2: Thesis overview

2 RESEARCH METHOD

This chapter will elaborate on the research method of this thesis. A simplified version of the research model is shown, followed by the methodology of the research in an illustration. In the first chapter the research approach is described. In the second chapter the validity and reliability of the research is discussed. The third chapter describes the contribution of this research and in the last chapter the research planning is described and depicted.

2.1 RESEARCH APPROACH

The research approach for this research can be divided into three major steps. In the three steps, different scientific methods are applied, because of the different purpose of each step. The first step is focused on gaining knowledge from the current literature, which will be obtained through scientific search engines. The second step will synthesize and validate the findings of the found literature in practice. In the third step, the findings of the previous two steps will be developed and validated within a project or case.

2.1.1 LITERATURE STUDY

The literature review analyses three domains, namely:

1. The current knowledge of VBHC;
2. The current knowledge of Process Mining in the healthcare sector;
3. Case studies which applied Process Mining techniques in VBHC.

Literature of the three mentioned domains will be obtained through databases with snowballing. The snowballing approach of Wholin (2014) will be followed and is shown in Figure 3. This research is also done by reading available literature and attending courses, meetings, and Masterclasses about Value-Based Health Care and Process Mining.

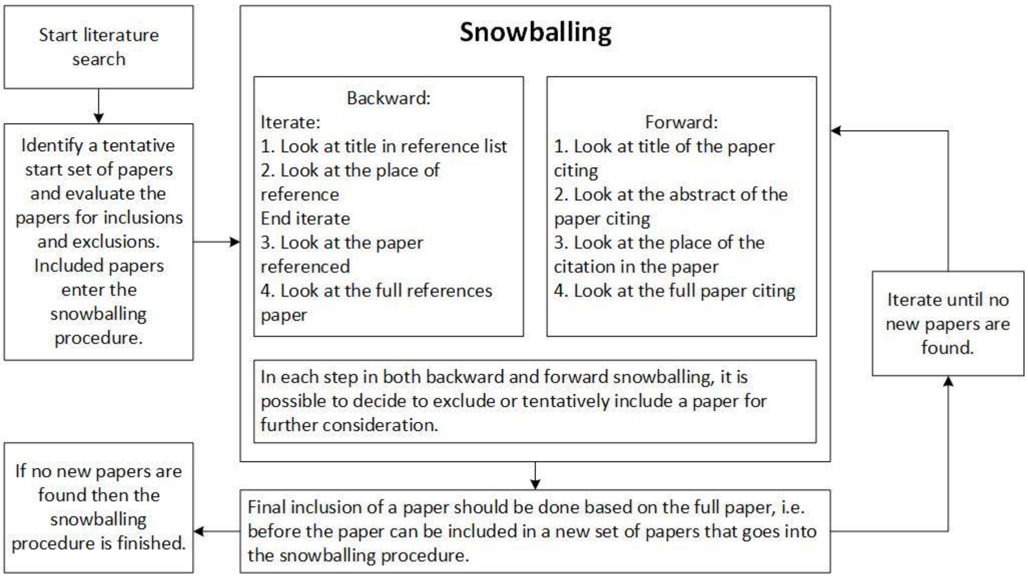


Figure 3: Snowballing Approach (Wholin, 2014)

The inclusions and exclusions criteria will be formulated after the collection of a set of tentative papers. An include criterion for example is freely availability of papers. The set of tentative papers are formed with papers gained from a masterclass of the organization The Decision Group, which is about Value-Based Health Care and Process Mining.

We will start by searching for papers in Google Scholar and PubMed by using the following terms: "*Value-Based Health Care*", "*Value in Health Care*", "*Process Mining in Health Care*" and "V*alue-Based Health Care in combination with Process Mining*". Papers will be first examined based on their title and abstract. Thereafter, with a set of strong informative papers the backwards and forwards snowballing technique will be performed as depicted in Figure 3.

For this research a method will be created using the theory of VBHC and Process Mining, which will be called Care Mining. This method will be created to be able to investigate whether a link between the two topics is possible. From VBHC the cost aspect will be used in order to generate more awareness about the costs of processes (Ronny S Mans, Aalst, & Vanwersch, 2015). Process Mining will use event logs for the discovery of processes and can be accompanied with time averages from practice.

2.1.2 CASE STUDIES

Besides a literature study, a case study is carried out. Object of study is the effect of Process Mining strengthening Value-Based Health Care in hospitals. This research needs actual data which is directly purchased from EHR systems from hospitals. Using real data from hospitals systems will lead to a more realistic and more complete representation of the current situation. Therefore, knowledge from managers and medical specialists is needed. The choice to use the first group is made because knowledge from managers about data of care activities and care processes will help to understand obtained data from EHR systems. The second group has practical knowledge about care activities and care processes which helps in understanding the current situation. Eventually, the aim is to increase the knowledge for managers and medical specialists about data of care activities and care processes in relation to costs. Knowledge from managers and medical specialists will be acquired by using interviews.

The case(s) that will be used are selected from cases acquired by Zuiver ICT. For the case study design the method of Yin (2009) is used which is depicted in Figure 4. The case study design consist of three main phases in which different activities are executed.

In the first phase “Define and Design” a theory is developed based on literature as is described in *2.1.1. Literature Review.* The theory is based on different approaches from literature. After theory building, appropriate case studies are selected based on the applicability of the built theory.

To provide qualitative information about the case studies a case report will be created in the second phase “Prepare, Collect, and Analyze.” A mixed method approach is chosen for this study, because it provides a means to answer different aspects of the complex research problem. Mixed methods research is formally defined as the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study (Johnson & Onwuegbuzie, 2004) and is stated to be helpful in field research (Edmondson & McManus, 2007). The qualitative side of the method is for the exploratory part of the research. It is used to gain an understanding of underlying reasons, opinions, and motivations of implementing Value-Based Health Care and Process Mining. It is also used to uncover similarities and differences in thought and opinions. The quantitative side of the method is used to quantify the research by generating numerical data which can be transformed to usable statistics, such as turnover times.

The pre-tests as well as post-tests consist of a direct observation and interviews. There is chosen to do both to combine the strengths of these methods. Direct observations are a means for the researchers to understand interpretation of the performed interviews about the performed activities and processes.

Finally, in the last phase “Analyze and Conclude” the theory is modified based on the conclusion drawn from the different case studies. The reasoning and modification of the topics VBHC and Process Mining are written down in the cross-case report and added to the master thesis.

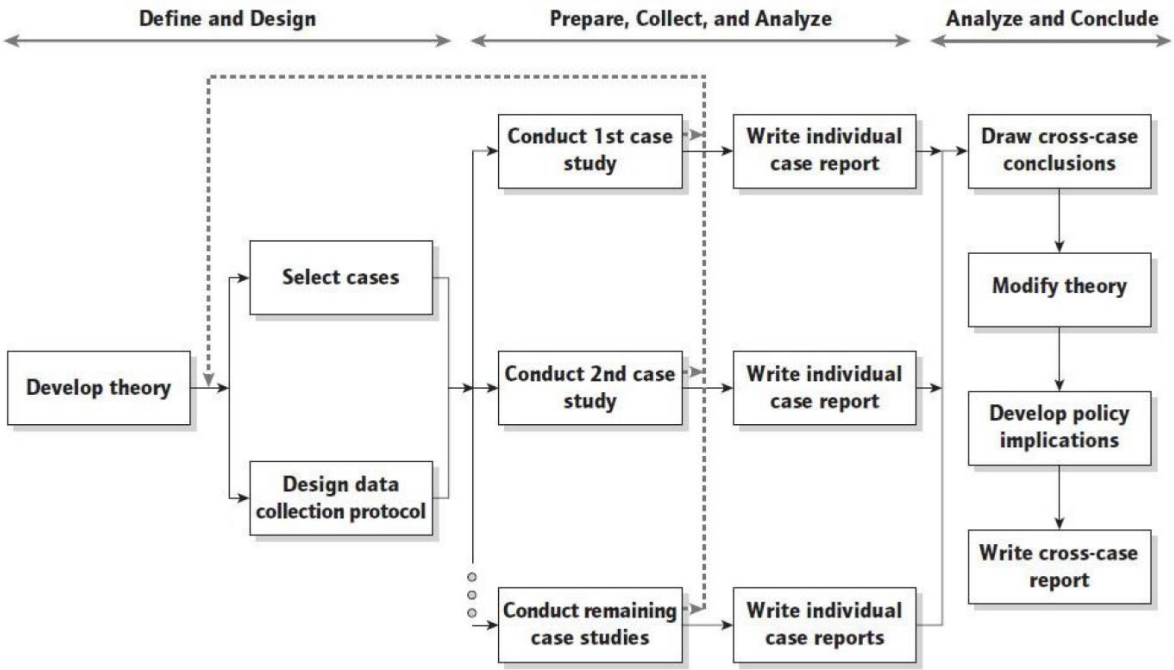


Figure 4: Case study approach (Yin, 2009)

2.1.3 DESIGN SCIENCE

This thesis is aimed at delivering a new artefact, which coincides with the definition of design science which is "devising artefacts to attain goals" (March & Smith, 1995). This artefact will be built to help cope the cost issues in hospitals. The artefact or method will be based on the findings of the literature study and will be designed with the existing theory of VBHC and Process Mining.

Design science secures that research is both relevant and effective. The framework of design science consists of the processes built and evaluate, which will be used in accordance to the specifications proposed by Hevner, March, Park and Ram (2004). In Figure 5 an application of this research context can be seen.

This research will be conducted in an environment consisting of health care providers and health personnel in hospitals. These parties are already trying to improve their efficiency and effectivity through productivity and financial management, but would like to enhance and increase it. Hospitals want to perform that by focusing on process improvements and creating insights into existing applications (which consists of patient related data). The knowledge base foundations consist of existing theory about VBHC and Process Mining. The environment and knowledge base provides business needs and applicable knowledge as inputs for the IS research. The methodology of TD-ABC will be used to build a method which analyses the applicability of Process Mining with VBHC. Finally, with the help of the evaluation, the method will be refined into a final version and re-evaluated by means of a repeatable research among health care managers.

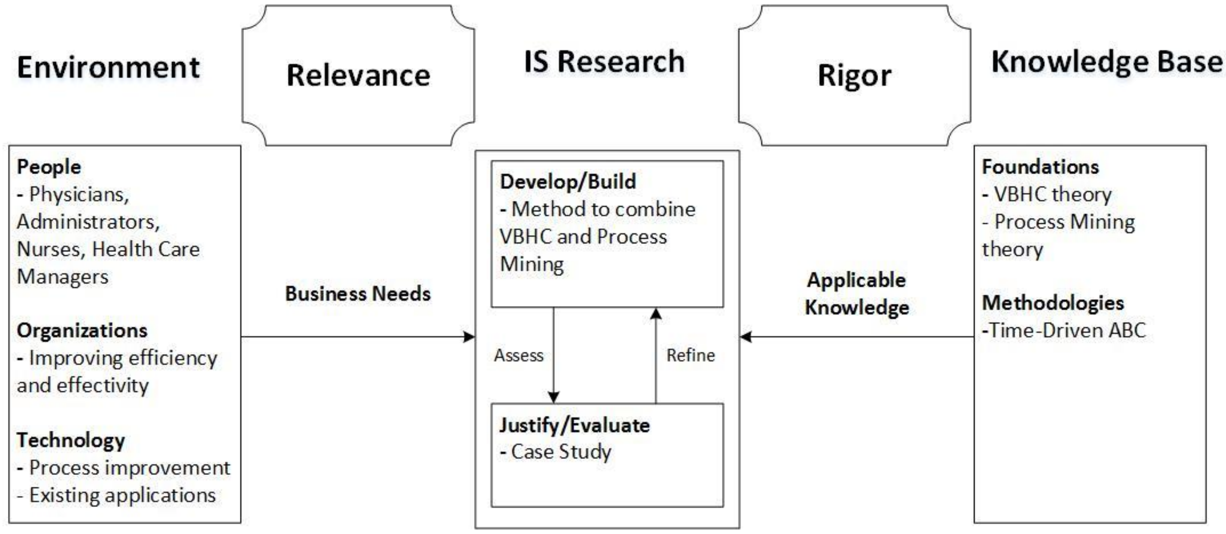


Figure 5: Design Science applied to this research

2.1.4 VALIDATION

The validation of the method, which will be created with literature study, is performed with the help of an internal person from the University. A person from Utrecht University will be arranged to conduct the method to see whether the results will be met with the results from the researcher of this thesis. For this validation data from the case studies of this thesis will be used. The results will then be validated with experts who were involved with the case studies.

2.2 VALIDITY

The research method chosen for this research has a qualitative character and therefore the validation of this research should take place. Yin (2009) presents criteria for judging the quality of the research design. The criteria should be taken beforehand, interpreted and applied by the researcher in the following way.

Construct validity

Construct validity is defined as *"[i]dentifying correct operational measures for the concepts being studied"* (Yin, 2009). To identify the correct operational measures the theory and case reports are peer reviewed by the end-users, but also by Zuiver ICT to check the quality.

The created method will be validated by the modeler of the method and by a medical specialist. The validation by the modeler will be performed using a scale construction (Hinkin, Tracey, & Enz, 1997). This construction is based on a Likert-scale with five points. The Likert-scale has proven that it creates variance, which is required when examining the link among different items (Lissitz and Green, 1975). The five-point scale consists of the following points, namely: easy, simple, normal, difficult, and challenging. Each of these scales are explained in the following part:

1. Easy – no problems were occurred, which could influence the performance time;
2. Simple – there were some minor problems, but had no influence to the performance time;
3. Normal – a number of problems occurred, but had no influence to the performance time;
4. Difficult – there were some minor problems and this influenced the performance time;
5. Challenging – several problems occurred and had a significantly effect on the performance time.

Internal Validity

According to Yin (2009) Internal Validity for case studies is: *"Seeking to establish a causal relationship, whereby certain conditions are believed to lead to other conditions, as distinguished from spurious relationships".* By documenting the starting point (for snowballing) from the literature review the base theory can be recreated. The data and analyses will be included in the results of this research to provide background information for the case reports. At the end the process, for creating the method, is presented using the results of the theories of VBHC and Process Mining.

Empirical Reliability

Yin (2009) defines Empirical Reliability as *"D[d]emonstrating that the operations of a study—such as the data collection procedures—can be repeated, with the same results".* This will be done by iteratively performing the snowballing techniques and allowing third parties to perform the composed method.

2.3 CONTRIBUTIONS

In this chapter both the scientific and social relevance of the project are described and discussed.

2.3.1 SCIENTIFIC RELEVANCE

This research investigates the current state-of-art in literature of VBHC with Process Mining. The current state-of-art is combined with the case studies. This research has not been conducted before in the healthcare sector and therefore helps academia understand the link between Process Mining and VBHC in practice more.

Also, with the theory of Process Mining the advantage and efficiency of creating process models with applications (such as; Disco or ProM) will be shown.

2.3.2 SOCIAL RELEVANCE

Social Relevance contains the usage of this research in the healthcare sector. The research provides the description of case studies for Process Mining in VBHC from a practical view. The current case studies highlight the techniques and technical aspects of VBHC and Process Mining and focus less on the added value. The technical aspect can be relevant for science, but less for the healthcare sector.

Also, a depiction will be created about the costs of care cycles.

2.4 RESEARCH PLANNING

In the following section, the planning procedure of this research is illustrated in a Process Delivery Diagram (PDD) (see Figure 6), which is created with the Method Engineering method of Brinkkemper (1996). It clears up how the chosen methodological approach contributes to reach the main objective of this research and thus to answer the sub-questions and main research question. To maintain an overview in the PDD, only the major milestones associated with the Master thesis. All activities will however be defined in the document.

2.4.1 PROJECT PLANNING

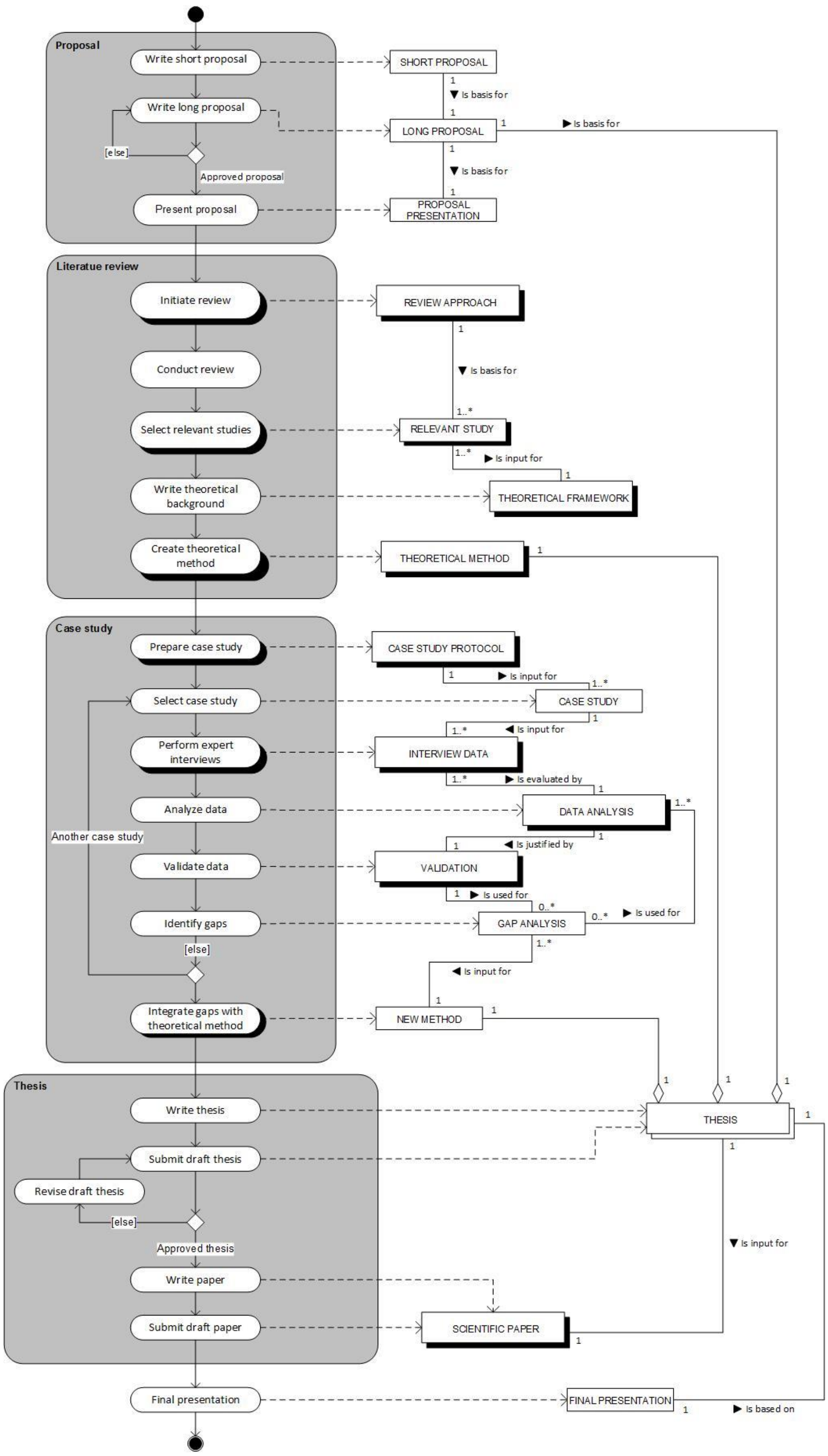


Figure 6: Project planning

3 VALUE-BASED HEALTH CARE

Value-Based Healthcare (VBHC) is based upon a theory which is developed by Michael Porter (2005). According to Porter VBHC is defined as the care outcomes which is spent per Euro (Porter, 2009). The costs include the full cost of the total chain of care around a patient (type). Other studies also refer to health results as Patient Reported Outcome (PRO) (Chen, Ou, & Hollis, 2013). Care Outcomes are the result of a whole care cycle chain (Chen et al., 2013). How they are defined precisely and which dimensions they cover, depends on the type of care. For example, in the care chain pregnancy and childbirth will eventually be about a healthy mother and a healthy child. In a chronic care example, besides clinical parameters, more emphasis will be on the quality of life.

Walter A. Shewart, a formal expert in the continuous quality improvement theory, stated that "[p]rice has no meaning without a measure of the quality being purchased" (Shewhart, 1980). Referring to the fact that quality and costs are related to each other. This can also be acknowledged when taking the notion of value into account. According to Brent C. James (1989) value has the following definition, which is: *"the combination of the quality of a product and the cost at which that level of quality is achieved."* Porter (2010) also mentioned that value is defined as: "the health outcomes achieved per dollar spent", where outcome is related to quality (Rubin, Pronovost, & Diette, 2001).

Quality of care can be approached in three aspects (Kunst et al., 1996; Kuenen, Mohr, Larsson, & Leeuwen, 2011). The first is focused on medical results of the conducted care. The results refer to outcomes indicators, which for example determines the morality of complications as a result of the performed treatments. The second aspect focusses on how care is being delivered. This could be about process- or structure indicators, which determines the methods and resources to carry out, and within which organizational context the treatment takes place. The third aspect focusses on the experiences of patients during a treatment. This could be determined with the use of surveys to ask about patient satisfaction, providing proper information, and practical issues (e.g. operation times). All the three manners are needed to approach quality management, but the emphasis should be placed on outcomes indicators. Hospital's ultimate goal is to raise and enhance health gains for patients. Unfortunately, The Netherlands and other countries in the world lack in outcome indicators (Kuenen et al., 2011).

In addition to quality and cost there is a third element of value of care. It is important to identify this component, because it has both the quality of care and it directly affects the costs. This element is the appropriateness of care. Appropriateness is defined as the degree to which the care for a patient is actually indicated (Kaplan, 2013).

The stakeholders that will benefit from value-based health care are health care providers, patients and suppliers. For providers, together with health plans, it means that they need to create competition on results and compete to each other on this level. Resulting, consumers or patients are more responsible and able to make better choices (Porter et al., 2013).

3.1 APPROACHING VALUE IN HEALTH CARE

The aspect *value* in health care can be approached from different starting points. Value can be approached from the side of the patient and from the side of the hospital. Both approaches are eventually directed towards the patient, which has been acknowledged by Kaplan and Porter by the following: *"The proper goal for any health care delivery system is to improve the value delivered to patients"* (Kaplan & Porter, 2011).

With the first approach, value is measured with information from the patients itself. They provide information to the hospital staff by means of online surveys, which is drafted with the help of medical staff. This information needs to be analyzed and translated into performed processes, to inquire the added value of implementing the information. The second approach aims at measuring value with data generated from hospital information systems. This approach focusses on the time and cost of performed processes. This information is collected from hospital's information systems, included with patient data. Eventually, both approaches will streamline the path of patients through the system, and select treatments that improve outcomes while eliminating services that do not.

3.2 MEASUREMENT OF VALUE

In the last few years, The Netherlands has widely published on value of care. When performing an Internet search with the terms 'value', 'value' and / or 'care' for example, publications are likely to be found from consulting firms such as Boer and Croon "To value creation in health care" (Boer & Croon, 2008), BCG's "Care for Value" (Kuenen et al., 2011) and KPMG's "Measuring Value of Healthcare Delivery" (KPMG, 2013). They are in any case interesting and valuable contributions, which have caused a significant development in health care (Dol, Raap, & Lageman, 2014; Hagen, 2015; van Deen, Esrailian, & Hommes, 2015). An important aspect is the conceptual meaning of the concept of value. Therefore, the following question needs to be answered: "How do you measure value of care?" An important aspect are the existing data sources and its strength lies mainly in combining these sources (Prodel, Augusto, Xie, Jouaneton, & Lamarsalle, 2015). When creating value for patients the most improved cost outcomes are the expected results.

Value of medical conditions can be determined with its results. These results are the health outcomes of patients per cost unit. When creating value-based healthcare the patient outcomes are in relative to its costs per Euro spent. In other words, value-based health care is the value of a particular treatment divided by its costs (Porter, 2005). The costs are the integral costs of a care chain of a particular patient or patient type. Value can inter alia reflect on the quality of healthcare and this will result in patients being better informed, better coordinated care, and expected positive care that is being rewarded. In order to create awareness and transparency about Value-Based Healthcare this abstract term should be made operationalized and measurable. This will be explained in the next sections.

Value-Based Health Care exists of three elements, namely: 1) health care outcomes (efficiency, safety and patient centeredness), 2) costs (the integral costs of a care chain), and appropriateness of care (van Deen et al., 2015; Porter, 2010). The appropriateness of care is the degree to which care for a patient is actually indicated.

In health care, there are predefined structures which defines a certain program for a medical condition. To perform the program resources are needed. With that program the treatment steps are defined and executed. Resulting, are the health outcomes from patients. Analyzing used resources and health services leads to a focus on efficient working. On the other hand, analyzing health services and health outcomes will lead to effectively evaluate care (see Figure 7), because the effect of a certain treatment is examined. Eventually when health outcomes improve costs will also reduce. This is the case due to the fact that health outcomes are giving a clear depiction of the current situation, when designed into a process model, and will show where costs can be saved.

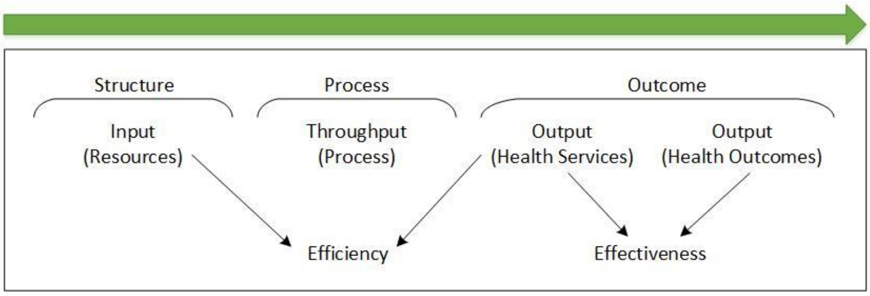


Figure 7: From structure to health outcomes (James, 1989)

Featured care is given conform certain guidelines which is predefined and researched indicated care. Along with this featured care appropriate care represents the area where care is desirable, thus known and measurable. The area where most problems occur are about indicated care, which is known beforehand, and practice variation, which differs per patient. The latter refers to situations which arises when for example complications emerges. In other words, the practice variation consists of desirable and undesirable care (see Figure 8). Undesirable care, which includes under treatment and overtreatment, is the reason that more expenses are incurred (Chen et al., 2013). This is the case due to undesirable care which is not predefined and creates unexpected situations. Because these undesirable processes are not modelled properly it is not visible where it has been carried out, how often it takes place, and what the additional costs are. All patients are different and thus these processes cannot be prevented, and it should not. What is needed is transparency about operationalized care and this should be made measurable.

Figure 8: Position of care guidelines, indicated care and practice variation

3.3 VBHC METHOD: TIME-DRIVEN ACTIVITY-BASED COSTING

Time-Driven Activity-Based Costing is a methodology which provides data about processes and its costs in order to redesign care (Hennrikus, Waters, Bae, Virk, & Shah, 2012; Lappi, Flath-Sporn, Heald, Kim, & Meara, 2015; Kaplan & Anderson, 2004). It helps in creating insight into the real made costs of a care cycle in order to redesign care (Kaplan & Porter, 2011). The methodology is designed by Kaplan and Anderson (2003) and identifies costs made in certain processes, redundancies in process steps, health outcomes that do not benefit a patient, and recognized unused capacity of people and equipment (Kaplan et al., 2014). TD-ABC provides data about processes and its costs in order to redesign care (Hennrikus, Waters, Bae, Virk, & Shah, 2012; Lappi, Flath-Sporn, Heald, Kim, & Meara, 2015; Kaplan & Anderson, 2004). It is measured by taking the costs of a resource of a treatment combined with the time a resource needs to execute an activity (Kaplan & Anderson, 2004). Kaplan and Anderson (2007) mentioned that TD-ABC: *"Can be used in any industry of company with complexity in customers, products, channels, segments, and processes and large amounts of people and capital expenditures."*

Hospitals are working with electronic health records (EHRs) which has resulted into generating massive datasets (Murdoch & Detsky, 2013). According to a survey by the American Hospital Association a duplication of the implementation of EHRs has occurred from 2009 to 2001 (Charles, Furukawa, & Hufstader, 2012), which means a duplication in datasets. The data from EHRs can be divided into quantitative data (e.g. laboratory values), qualitative data (e.g. text-based documents), and transactional data (e.g. a record of medication delivery). Unfortunately, this data is not being treated as a central asset in order to improve efficiency and effectiveness, but rather as a by-product of health care delivery (Murdoch & Detsky, 2013).

The TD-ABC method can work with complex data, variable processes, and produces less data. Therefore, this will result in a more achievable method for hospital data (Hilsenrath, Eakin, & Fischer, 2015). This method uses clear steps in creating an overview about costs (in time and finances), which will be explained later in this section. When creating an overview about costs the direct and supportive capacities will be predominantly taken, what means that the costs around patients will be used. According to Cooper and Kaplan (1999) the goal is to be "approximately right" (which could be between 5% to 10% of actual numbers) and not "precisely wrong", because emphasis should not be much on intensive time consuming studies. Over the time, TD-ABC will reveal any error which could have been made.

Current cost measurement has led to the following issues (Kaplan & Porter, 2011):

- Aggregating and analysing costs at procedure, specialty or service department level, instead of on an individual patient level.
- Cutting payor reimbursement, which does nothing to reduce the actual costs.
- Inability to link cost to process improvements or health outcomes.
- Cross-subsidies across services, which lead to deformations in the supply and efficiency of care.

To cope with these issues Time-Driven Activity-Based Costing will be introduced.

3.3.1 TD-ABC TECHNIQUE IN STEPS

TD-ABC requires two parameters to be estimated, namely: 1) the cost unit of capacity, and 2) the time required to perform an activity (Kaplan & Anderson, 2003).

3.3.1.1 ESTIMATING THE COST UNIT OF CAPACITY

The first parameter is about the practical amount of resources that is available per department (i.e. creating an activity directory). This phase is aimed at determining the amount of capacity per department. With the first phase, an analysis per minute needs to be made to assess the availability of personnel and machines. The time in minutes or hours is achieved by interviewing the concerning medical staff who performs certain processes. The cost per department are obtained from existing rates or an estimation is made by interviewing the concerning medical staff.

Additionally, the practical availability of the hospital staff is determined. The practical availability is without consultation, breaks, arrival, departure, communication, and training time. Kaplan & Anderson (2004) adopts a practical availability of 80% for staff and 85% for machines. In Dutch literature it is referred to as "normal capacity" (Mutze & Ierland, 2007).

3.3.1.2 ESTIMATING THE TIME REQUIRED TO PERFORM AN ACTIVITY

In the second phase, the time for the execution of an activity should be determined (i.e. the output). The second step is determining the amount of time it cost to conduct certain activities, which is executed by means of interviews (optionally surveys) or direct observation.

Finally, the cost price per activity is calculated by multiplying the practical availability and costs per activity (Kaplan, 2013).

The aforementioned costs are directed towards a department level, but can be narrowed down to a patient level cost-determination. Figure 9 depicts detailed steps about how to determine total costs for patients in a care cycle.

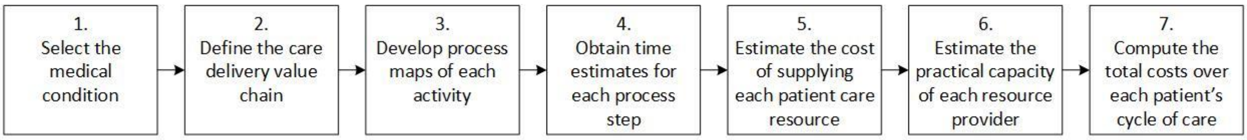


Figure 9: TD-ABC steps for patient centered cost measurements (Kaplan & Porter, 2011)

1. **Select the medical condition**

The first step requires to define a scope by selecting a medical condition, for which a cost estimation will be made per patient. In most cases, a demand originated from the hospital allows to select a medical condition. In other cases, interviews with managers will indicate which medical condition has a priority to be analyzed.

Scoping a medical condition will be done by defining the start and end process of a selected medical condition of a certain patient population. Estimating the start- and endpoint is possible by taking interviews. These interviews require managers or medical specialists to explain how and when a patient enters and leaves the hospital.

2. **Define the Care Delivery Value Chain**

The second step is about specifying the Care Delivery Value Chain (CDVC). The Value Chain is a concept to capture the strategic relevant activities of an organization. It was developed by Professor Michael E. Porter and can be traced back to his 1985 Harvard Business Review article, "*How Information Gives You Competitive Advantage"* (Porter, 1985). It is frequently used to visualize a "Business Unit" in the industry and it breaks the business unit's value creation process down into *primary activities* and *support activities*. The elaboration of a value chain often helps the management of an organization to recognize issues where value and strategic positioning can be optimized.

Porter argues that "competitive advantages arise from choices in the Value Chain". As an introduction to this topic, Porter's standard Value Chain is described briefly. Then, the Care Delivery Value Chain is described in more details. Porters basic Value Chain can also be used by a hospital or clinic to analyze general activities, which are not directly associated with care delivery. Such general primary activities would be "Inbound Logistics", "Operations", "Outbound Logistics", "Marketing and Sales", and "After Sales Services". Support activities in the basic Value Chain are "Firm Infrastructure" (e.g. financing, planning, investor relations), "Human Resource Management" (e.g. Recruiting, Training, Compensation System), "Technology Development" (e.g. Product Design, testing, process design, material research, market research), and "Procurement" (e.g. components, machinery, advertising, services) (see Figure 10).

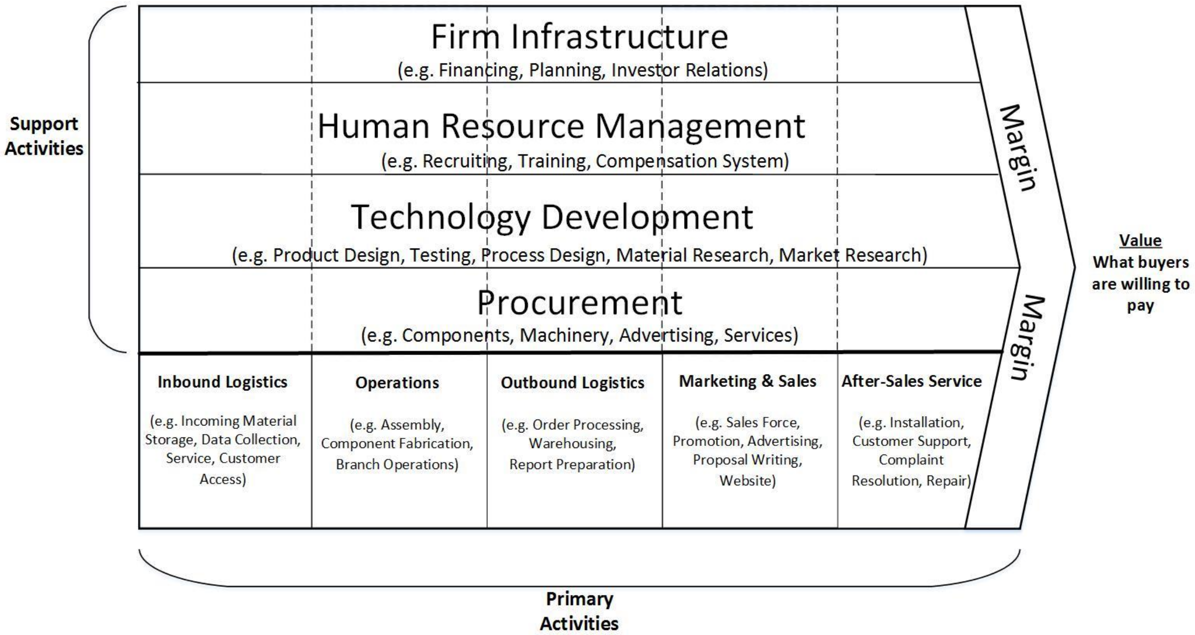


Figure 10: Original Value Chain (Porter & Millar, 1985)

In contrast to the classic Value Chain, the Care Delivery Value Chain focuses on health care services. The activities in the CDVC are broadly defined to include all important aspects of a certain medical condition. By using the CDVC an overview will be created of a patients full care cycle (see Figure 122), which helps in improving most processes and lean health care (Porter & Teisberg, 2006). The latter aims for a continuous flow of patients and the care provider (Kaplan & Porter, 2011). The value chain can be approached as a tool in order to map all the involved activities or processes used in delivering care. The activity steps are obtained by interviewing managers or medical specialists. The following sections describe the principles of the Care Delivery Value Chain, which could become a standard practice to analyze the strategic positioning of a clinic or hospital.

The full cycle of care of a defined medical condition is divided in the direct value added *primary activities*, and indirect value-added *support activities*. Primary activities are "Monitoring and Preventing", "Diagnosing", "Preparing", "Intervening", "Recovering", and "Monitoring and Managing". Support activities are "Informing", "Measuring", and "Accessing" (see Figure 11).

Each layer can be elaborated on each process of a typical path of patient care. The processes on the path of patient care is structured from *preventing* to *managing* because in general it reflects how most doctors are trained, it is traditionally the hospital-centric nature of the system, and the local custom of the provider organizations (Rubin et al., 2001; Porter, 2008). Also, it reflects a certain way of working in health care where doctors are expected to handle everything, and each facility is expected to be able to provide every service (Rubin et al., 2001). Since providers cannot predict what each patient will need beforehand, the delivery of care is organized around general purpose units in order to handle any issue in each discipline (Porter, 2008).

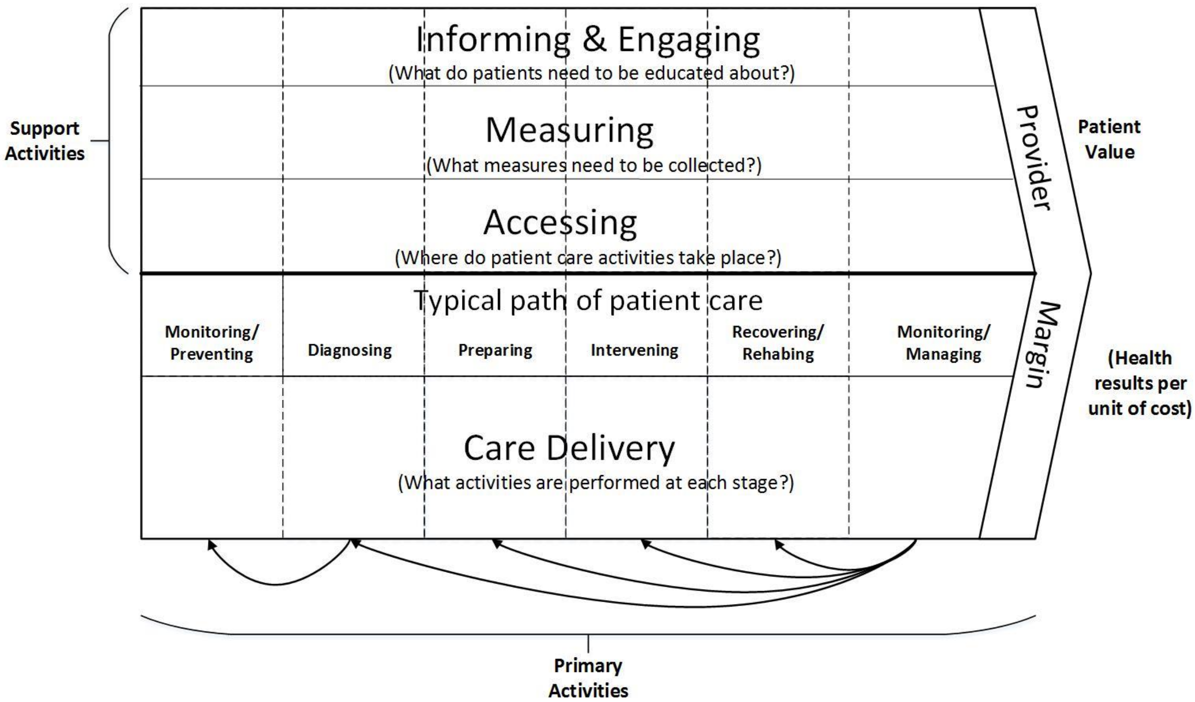


Figure 11: The Care Delivery Value Chain format (Porter & Teisberg, 2006)

The aim of every activity should be to optimize patient value. Porter and Teisberg (2006) argue that patient value would be to get the best health results per unit of cost. The CDVC of Porter (2008) describes that defining the activities in the care cycle leads to a margin of the whole cycle. This study does not focus on the financial aspects of VBHC.

PRIMARY ACTIVITIES

Primary activities are direct value added. They represent the full cycle of care for the patient, from preventing, diagnosing, over preparing, intervening, to recovering, monitoring, and managing. These activities are explained in the remainder of this section.

1. Monitoring and Preventing

For any medical condition, the path of patient care starts with monitoring and preventing. Ideally, this leads to an early detection of a disease. The Value Chain raises awareness of interceptions of prevention activities. It could be a great chance for a hospital to cooperate with other health authorities, schools, patient organizations, and employers of risk groups—and make links to medical specialists in the hospital. The links to referring physicians should be managed carefully.

“Monitoring and preventing” shows opportunities of optimizing the volume of patient intakes. How do competing hospitals in the region position themselves to get patients? Are they just there and wait until patients arrive? How can patient value be optimized with monitoring and preventing activities?

If patient value is increased, the marketing of the hospital is justified and is a strategy to position the hospital as center of excellence. The involvement in prevention is a statement of delivering excellent care.

2. Diagnosing

Early detection of diseases often helps to reduce their impact on the patient’s health. Hence, early referral to specialists with suitable equipment can improve patient value. However, as patient value is defined as health results per unit of cost, specialized equipment should be allocated in a reasonable manner per region to avoid the effect of offer induced demand (described later in this paper).

In diagnosing, medical and family history should be checked routinely, and, if appropriate, dedicated advanced testing initiated. The links for communication and consultation with other specialists (e.g. data integration with modern information technology) should be managed. Finally, the treatment plan is formulated.

3. Preparing

To perform a successful intervention, the patient must get a procedure-specific preparation. A good preparation should optimize the patient’s convenience during the stay in the hospital. High quality orientation, proper information, and transparency increases trust which is an important element for successful health care services.

Ideally, the patient has a contact person, such as a case manager who is accompanying him/her trough the stay in the hospital. It may also be an advantage to conduct a needs analysis before an intervention. Sometimes very small things that could easily be changed have a big impact on patient comfort and patient value.

4 Intervening

Depending on the disease, intervention can occur with drugs and/or with a procedure (e.g. surgery). The most suitable technology should be applied by specialists. This specialization should push forward the learning curve and finally increase quality.

5. Recovering and Rehabilitating

Patient value is not only delivered by the intervention. Many (still) resource-oriented hospitals neglect the importance of recovering and rehabilitating. As in modern business, “after sales service” is crucial for a company’s reputation, success, and even survival. Hospitals should have the chance to position themselves with a better recovering and rehabilitating environment. It does not have to be more expensive, but better aligned with the patient’s expectations.

Re-hospitalizations due to neglected recovering after an intervention should be avoided. Again, a suitable needs analysis should be conducted with the patient. The drug regimen should be fine-tuned. There should be an incentive to find best regimen for patient, not for the doctor's income. Finally, creating patient value has a lot to do with patient orientation. Cost accounting is also important, but the patient should feel like a human and not like a number.

6. Monitoring and Managing

The path ends with monitoring and managing a condition, which could be ongoing for certain condition. For example, patients with cancer who were successfully treated often need to be managed for long periods of time. Since cancer could, become a chronic condition for many patients.

In monitoring and managing, medical specialists and administrative workers should keep a track record of their patients. To reach excellence, health care professionals should always learn and improve. The patient's compliance to the drug regimen should be followed-up.

SUPPORT ACTIVITIES

The support activities are not direct value-added, but are also an essential part of the Value Chain. Like the primary activities, all support activities are supposed to be aligned to the stakeholders' needs and create patient value. Especially in health care, communications are very important. Support activities visualized in the Care Delivery Value Chain are analyzed to discover synergies.

1. Informing

Informing is about communications with the patient. It starts with risk factor and lifestyle counseling under "monitoring and preventing". While diagnosing, the implications must be explained. Before the intervention, the patient must be educated on procedures. Finally, the patient's compliance to the treatment should be followed-up. Good health care professionals are trained in adequately informing patients.

However, daily routine may lead them to shorten this information. Informing is an important support activity which could be professionalized with a case manager who accompany patients during their stay in the hospital and inform patients before and after their intervention.

2. Measuring

Suitable technology should be available to correctly measure the patients' condition. Good physicians may choose where they want to work also concerning the state-of-the-art equipment available to perform their work.

Though, the most expensive equipment may even reduce "patient value" (health results per unit of cost) when unimportant things are measured. It should be taken care of the critical mass of measurement equipment of what is needed in the target population of the hospital.

3. Accessing

What may sound as a matter of course is sometimes not obvious in hospitals. Physical access to hospitals can be managed and enhanced. Emergency services must be accessible, or, if not available on weekends, patients must find access information easily.

Managing accessing means to co-ordinate office, lab, and hospital visits, as well as providing access to information via telephone or internet interactions.

An example of a full care cycle defined in a CDVC is depicted in Figure 12.

TYPICAL PATH OF PATIENT CARE	Monitoring/ Preventing	Diagnosing	Preparing	Intervening	Recovering/ Rehabbing	Monitoring/Managing
INFORMING AND ENGAGING What do patients need to be educated about?	• Importance of exercise, weight reduction, proper nutrition	• Meaning of diagnosis • Prognosis (short- and long-term outcomes) • Drawbacks and benefits of surgery	• Setting expectations • Importance of nutrition, weight loss, vaccinations • Home preparation	• Expectations for recovery • Importance of rehab • Post-surgery risk factors	• Importance of rehab adherence • Longitudinal care plan	• Importance of exercise, maintaining healthy • weight
MEASURING What measures need to be collected?	• Joint-specific symptoms and function (e.g., WOMAC scale) • Overall health (e.g., SF-12 scale)	• Loss of cartilage • Change in subchondral bone • Joint-specific symptoms and function • Overall health	• Baseline health status • Fitness for surgery (e.g., ASA score)	• Blood loss • Operative time • Complications	• Infections • Joint-specific symptoms and function • Inpatient length of stay • Ability to return to normal activities	• Joint-specific symptoms and function • Weight gain or loss • Missed work • Overall health
ACCESSING Where do patient care activities take place?	• PCP office • Health club • Physical therapy clinic	• Specialty office • Imaging facility	• Specialty office • Pre-op evaluation center	• Operating room • Recovery room • Orthopedic floor at hospital or specialty surgery center	• Nursing facility • Rehab facility • Physical therapy clinic • Home	• Specialty office • Primary care office • Health club
CARE DELIVERY What activities are performed at each stage?	**Monitor** • Conduct PCP exam • Refer to specialists, if necessary **Prevent** • Prescribe anti-inflammatory medicines • Recommend exercise regimen • Set weight loss targets	**Imaging** • Perform and evaluate MRI and x-ray - Assess cartilage loss - Assess bone alterations **Clinical Evaluation** • Review history and imaging • Perform physical exam • Recommend treatment plan (surgery or other options)	**Overall Prep** • Conduct home assessment • Monitor weight loss **Surgical Prep** • Perform cardiology, pulmonary evaluations • Run blood labs • Conduct exam	**Anesthesia** • Administer anesthesia (general, epidural, or regional) **Surgical Procedure** • Determine approach (e.g., minimally invasive) • Insert device • Cement joint **Pain Management** • Prescribe preemptive multimodal pain meds	**Surgical** • Immediate return to OR for manipulation, if necessary **Medial** • Monitor coagulation **Living** • Provide daily living support (showering, dressing) • Track risk indicators (fever, swelling, other) **Physical Therapy** • Daily or twice daily PT sessions	**Monitor** • Consult regularly with patient **Manage** • Prescribe prophy-lactic antibiotics when needed • Set long-term exercise plan • Revise joint, if necessary

Figure 12: Example of a CDVC from a Severe Knee Osteoarthritis Replacement (Kaplan & Porter, 2011)

3. Develop process maps of each activity

In this step, all the activities a patient follows in a care cycle will be modelled in a detailed process map. This process model includes capacity-supplying resources (which means personnel, facilities, and equipment) for each activity in the care cycle. An example of a process model is depicted in Figure 13. Often, process models are modelled by using the Business Process Modelling Notation (BPMN) or UML activity diagrams (Goedertier & Vanthienen, 2006). BPMN is a standard which models business processes in a graphical manner by using a business process diagram. The latter is done to clarify the management of business processes and focusses on the fact that it is both understandable for technical users and non-technical users (Weske, Hofstede, & van der Aalst, 2003; White, 2004). The process model can be created by using several approaches. One of them is interviews or meetings. Another method which can generate a process model is Process Mining. This method or theory will be explained in the next section.

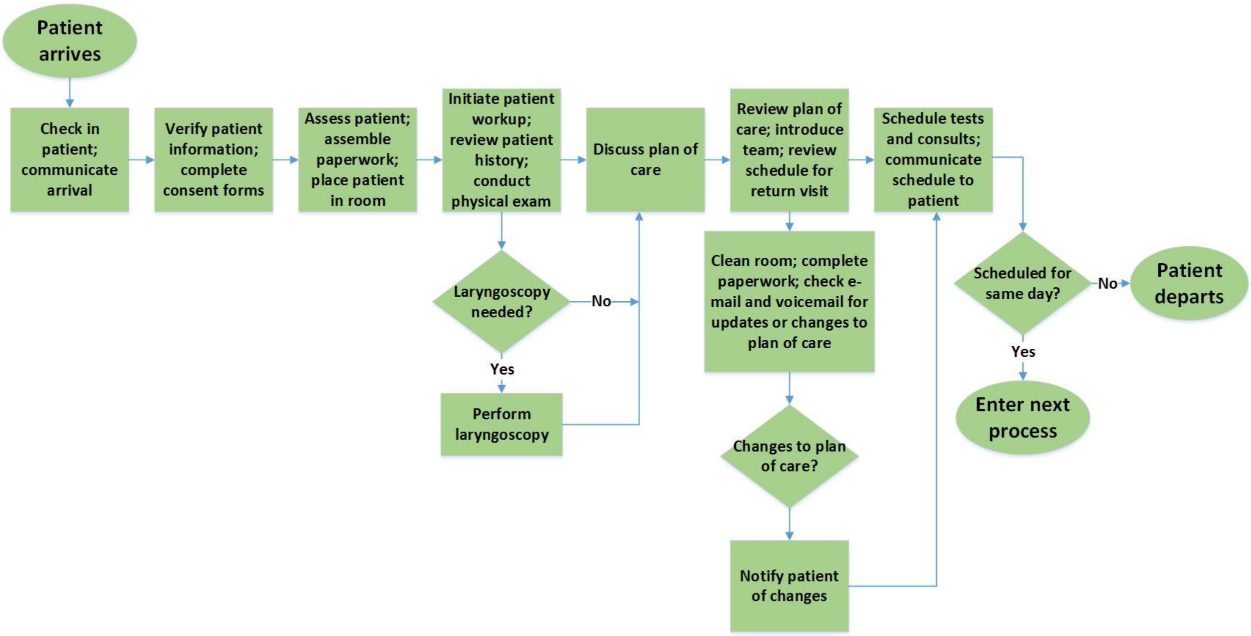


Figure 13: Example of a New-Patient Process Model (Kaplan & Porter, 2011)

4. Obtain estimates for each process step

In this step the time each providers or resource spends with a patient is estimated. Every set of resource will have a total amount of time spent. When there are different resources, they are separately estimated. Take for example patient X. This patient will visit a hospital for a first meeting, so the process name could be “First meeting”. Patient X goes to the hospital and meets three hospital employees for this appointment, namely: administrative personnel, a nurse, and a physician. The time estimation could be obtained with interviews or direct observations. With that it can be noticed that the patient spends 18 minutes (0.3 hours) with the administrator, 24 minutes (0.4 hours) with the nurse, and 9 minutes (0.15 hours) with the physician.

5. Estimate the cost of supplying each patient care resource

This research is focusing on costs in time and not in finances. Therefore, this step is optional or applicable in a further research.

This step concentrates on the costs each provider or resource spends on patients. All costs are related to the provision of hospital personnel (i.e. doctor, nurse and or administrator). The costs include: salaries, payroll taxes, possibly secondary benefits (health insurance and pensions). Furthermore, there are other related means to ensure that an expert is available to a patient. Namely: employee supervision (based on number of people under one supervisor), space (office of each employee, based on occupation of space and rental expenses), equipment, information technology (based on individual use of computers and communication tools and services) telecommunication of each employee in the normal performance of the work.

These data can be inter alia found in IT systems, the budgeting system, etc. The total of this data will be used to calculate each resource capacity cost rate. To obtain the costs made by a health care personnel the time spent for patient related work needs to be multiplied by the salary. With this method, one can notice that the remaining amount can be allocated to trainings, meetings, etc. The next step is to determine supplies which is needed to perform patient care (i.e. occupancy of space).

6. Estimate the practical capacity of each resource provider

The practical capacity is about the availability of an expert for patient related work. Based on one year, one can start with 365 days. These days will be deducted every day on which an employee is unavailable for work. Consider the following: weekends, days off, holidays and sick leave. Furthermore, the available times on one day will be calculated, which then will be multiplied by the total days per month (to calculate the total number of available hours per month). The hours that someone has no patient contact will be subtracted afterwards. For example: research and education.

7. Compute the total costs over each patient's cycle of care

Till now all the data which is needed to calculate the resource capacity cost rate has been determined. In this final step the total costs of resources (step 5) needs to be divided by its practical capacity (step 6). To compute the total costs capacity per patient per care cycle the capacity cost rate needs to be multiplied for each resource by the amount of time of patient care (Step 4). Lastly, the results of both calculations must be summed up. In Appendix 12.1 an example of this calculation can be found.

4 PROCESS MINING

Process Mining is a technique for process management that is used for the analysis of business processes, what is based on event logs (van der Aalst et al., 2012). An event log is a range of attributes that are linked to one case and preferable refers to an activity (van der Aalst, 2011). Besides, Process Mining creates overviews containing different disciplines, locations, etc. Therefore, it is also used to understand deviations, inspect bottlenecks, and monitor organizational behavior (Mans et al., 2015). Knowledge from event logs are extracted by data-mining algorithms and are obtained from information systems (e.g. databases, transaction logs, audit trails, etc.). Organizations use information systems to support the implementation of their business (van der Aalst et al., 2007). Examples of information systems that support operational processes are: Workflow Management Systems (WFMS): manages workflows with documents and/or information carried out by humans and/or systematic operations in an automated system, and Enterprise Resource Planning (ERP): integration and automation of business processes, e.g. SAP. Quality requirements are often set to these information systems in the context of efficiency and reliability (Bhattacherjee, 2012). The quality of efficiency is related to the following: flexibility- this quality relates to the adaptability and extensibility of information systems; maintainability- this is about the ability to quickly recover errors; integrability- this involves dividing information into subsystems (Monroy, Nasiri, & Peláez, 2014). The quality assurance must deal with the integrity (internal control) and auditability of data in information systems. By analyzing the actual course of the business in information systems, the quality of reliability and efficiency of the information can be analyzed using Process Mining (Mueller-Wickop & Schultz, 2013).

The aim of Process Mining is achieving efficiency and understanding processes by presenting techniques and tools. This is done for discovering processes, control data, organizational, and general structures from the event logs (Medeiros, Karla, van der Aalst, & Pedrinaci, 2008). In the industry this is also known as Automated Business Process Discovery (ABPD) (Gartner, 2016).

4.1 PROCESS MINING TECHNIQUE

The techniques of Process Mining allow to work with huge amounts of event data, which for example includes patient-related data, and translates this into organized descriptions of real activity paths. An advantage of these created models is that they are descriptive, in contrast with other modelling efforts (which are typically prescriptive) (Strausbert, Bilir, Waydhas, & Ruchholtz, 2003; Lenz & Reichert, 2007). The techniques of Process Mining are frequently used when there is a lack in obtaining formal descriptions of processes with other approaches, or when there is a lack in the quality of existing documentations (Kirchmer, Laengle, & Masias, 2013). An example where the techniques can be used to discover various models, describe processes, organizations, and different products, are the audit trails (or event sequence, which is the full care cycle of a specific patient) of a WFMS, the transaction logs of an ERP system, and the electronic patient records from a hospital system. These event logs can also be used for a comparison between event logs and process models to see whether the observed reality corresponds to a prescriptive or descriptive model.

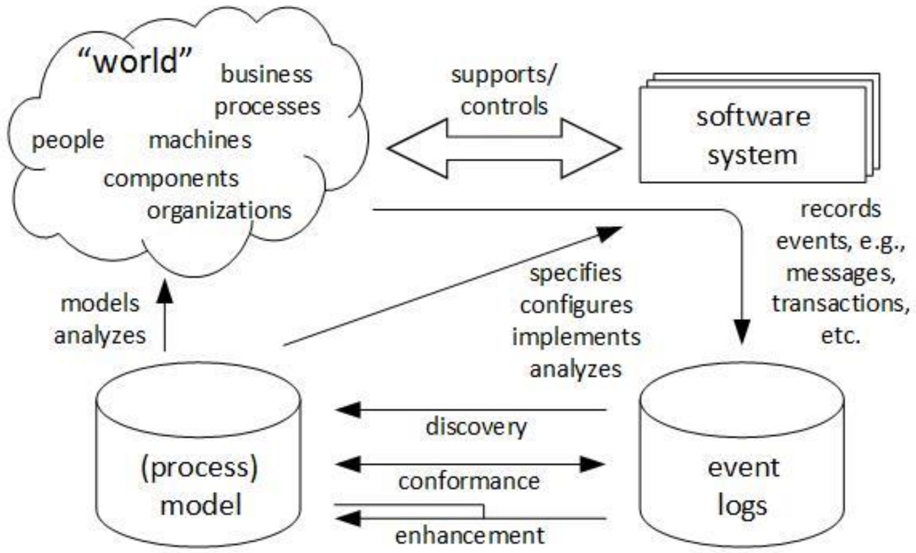


Figure 14: Process Mining definition (van der Aalst et al., 2012)

The Process Mining techniques can be divided into three main types, namely: discovery, conformance analysis, and enhancement (van der Aalst et al., 2012). The types are based on the existence of a model and the usage of it (see Figure 14).

Discovery: In the discovery analysis there is no priori model. Therefore, models are created which are based on event logs. The discovery of a process model can be done based on low-level events. There are several techniques which automatically can construct process models (e.g. Petri net expressions) and are based on certain event logs (van der Aalst, Weijters, & Maruster, 2004; Agrawal, Gunopulos, & Leymann, 1998; Cook & Wolf, 1998). Lately, Process Mining research also initiated to focus on other perspectives, such as: data, resources, time, etc. An example is the technique which is described by Aalst, Reijers and Song (2005) and can be used to create a social network (van der Aalst, de Beer, & van Dongen, 2005).

Conformance checking: In the conformance checking a priori model exists. The model is used to compare event logs and analyses differences between the log (i.e. the reality) and the model. An example, when this could be applied, is in the "four-eyes" principle. The conformance checking could be used to detect certain deviations to enrich the model. For example, expanding a process model with performance data. Another example is mining decisions which is described by Rozinat and Aalst (2005) and takes the priori process model and uses it to analyze each choice in the process model. The event log is consulted for each choice to analyze what kind of information is available at every moment a choice is made. To check which data elements could influence a choice, data mining techniques are used.

Enhancement: In the enhancement analysis a priori model also exists. The objective of enhancement is to extend or improve the model with a new aspect or perspective, which is in contrast to conformance checking where the alignment between reality and a process model are measured (van der Aalst, 2011). An example is extending a process model with performance data (i.e. using annotations, like coloring parts, to detect differences in the process model).

4.2 PROCESS MINING TOOLS

There are two types of Process Mining tools developed on the market, namely: open source and commercial.

To use Process Mining tools, the data in datasets need to have certain demands. All the Process Mining tools make use of event logs. About the content of event logs the Process Mining Manifesto states the following:

"[that all] Process Mining techniques assume that it is possible to sequentially record events such that each event refers to an activity (i.e., a well-defined step in some process) and is related to a particular case (i.e., a process instance). Event logs may store additional information about events. In fact, whenever possible, Process Mining techniques use extra information such as the resource (i.e., person or device) executing or initiating the activity, the timestamp of the event, or data elements recorded with the event (e.g., the size of an order)" (van der Aalst, et al., 2012).

This statement describes that an event log must contain a case identifier, timestamp and activity name.

For this research, two tools are used for executing the Process Mining analyses: (1) Fluxicon Disco, and (2) The Process Mining Toolkit (ProM), and Celonis.

ProM (Process Mining) is toolset maintained by Eindhoven University of Technology, which collects prototypes developed in several research projects. The project is open source and aims largely the academic and research group. The plugins added on demand enables to solve complex process exploration. ProM imports event logs compliant with the MXML or XES formats and can load process model definitions in different standards. Some of the main features of ProM are: discovering the control-flow perspective of a process, social network analysis, analyzing the resource and performance perspective of a process, discovering events based on decision rules and conformance checking with a variety of algorithms. ProM provides several export formats such as CSV and PNG.

Disco is a commercial process mining tool developed by Fluxicon, which can run on top of Windows or Mac Desktops. It supports a wide range of event log import formats including CSV, MS Excel, MXML, XES, FXL Disco Logs and DSC Disco project files. Some of the features include automated process discovery, animation of process maps, event log filtering with various parameters, project management and detailed statistics.

Celonis is a commercial company that provides yet another software as a service process mining tool. Among the features for this tool include Automated Integration of source data, real-time surveillance of all business transactions, execution of process analyses, various filtering mechanisms and process reporting.

4.3 PROCESS MINING METHOD

Little is known about the Process Mining method in literature. In 2015, only five papers extensively describe a process (or make modifications) for a Process Mining project. Bozkaya et al. (2009) are the first in the literature to describe a method of Process Mining. In this paper a case study is performed and an approach of six steps is defined. In Figure 15 the method proposed in the paper can be seen.

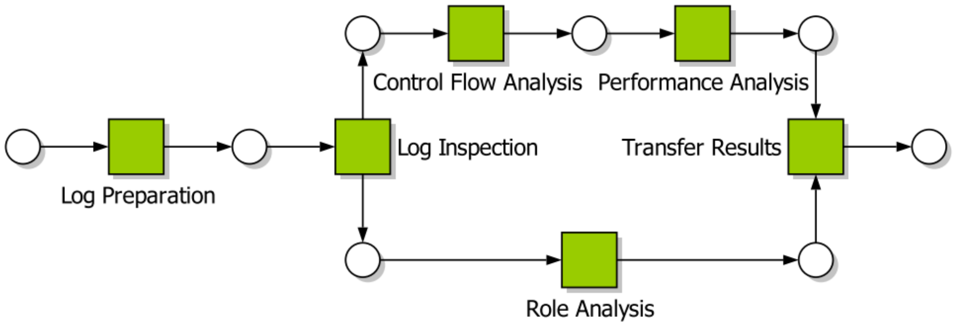


Figure 15: Process Mining method (Bozkaya, Gabriels, & Werf, 2009)

Bozkaya, Gabriels and Werf (2009) describe that Process Mining gives insight about processes based on the actual current situation and argues how analyses can be made from system's data, which will be followed in this research. In the paper the following phases are being acknowledged: (1) log preparation, (2) log inspection, (3) control flow analysis, (4) performance analysis, and (5) role analysis. Afterwards the results needs to be transferred to the client in order to validate the achieved results (Bozkaya et al., 2009). This methodology is applied for process diagnostics within a short period. The results of this approach lead to different outcomes, depending on each case and dataset.

For this research, the objective is to create models and analyses for process diagnosis with Process Mining. It allows Process Mining to cultivate on a practical level, rather than only focusing on the theoretical processes (Prodel et al., 2015). Applying this technique in different organizations would obtain a competitive advantage and therefore increase the maturity level of it (van der Aalst et al., 2007).

Log Preparation

The Process Mining method of Bozkaya et al. (2009) begins with the step Log Preparation. In the step the log file is being prepared for the importation in a Process Mining Tool(s), with the use of different steps. The log files are exported from (different) information systems.

Log Inspection

A Log Inspection is performed to create a first impression of the extracted log file. The focus relies on the number of events, cases and activities in the log file. This information can be used in a chart for certain analysis and to create a provisional process model. This step is important when analyzing organizations with large and complex processes.

Control flow analysis

Control flow analysis checks the accordance of the processes in the created process model to the reality. This is done by comparing and validating the event logs from information systems to predefined process models. Figure 16 shows an example of a control-flow, created with Petri nets, in the health care setting and describes a pathway, which could be from a specific patient or a patient group.

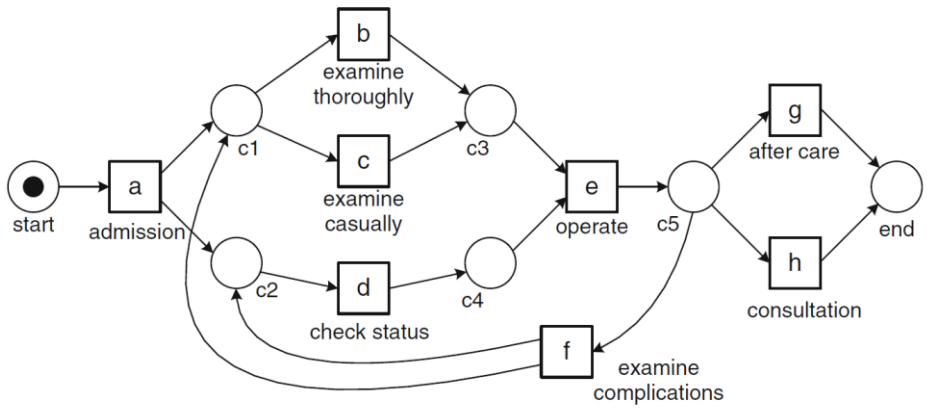


Figure 16: Control-flow of process model (Mans et al., 2015)

Performance Analysis

Performance analysis checks the throughput times of cases. Also, the throughput times of activities between each other could be checked.

In the case of a hospital's log file, several activities are present. Analyzing large numbers of activities can be difficult, but a dotted chart can be created and used. This chart generates a graphic visualization of the activities and makes it less complicated to analyze activities and cases. Generating such a chart can be done with the use of a Process Mining Tool, such as ProM (Bozkaya et al., 2009).

Role Analysis

As can be seen in Figure 15 the role analysis of Bozkaya et al. (2009) is performed parallel to control flow- and performance analysis. In role analysis, the resources are related to the activities in the process model. Thus, this creates an overview of who performed what activity in what amount of time.

Transfer Results

After several analyses are created, the results are handed over to the customer. This could be done in a (multidisciplinary) meeting with the use of a presentation. Also, a document could be communicated. The performed analysis could lead to (more) discussion and questions about the processes.

The different analyses are communicated to the customer. The medium of the transfer varies from a document to a presentation. The analyses can lead to discussions about the process with the customer. Bozkaya et al. (2009) shows the importance of these steps. Certain conclusions of the analyses can be recognized by the customer.

4.4 PROCESS MINING CASES

The Process Mining cases in hospitals are limited and this can also be seen by the limited number of care organizations (Caron et al., 2014), which deliver Process Mining as a service, to realize an improvement in their performed care processes. Derived information of data provides new questions about the current situation. With the help of Process Mining-software answers can be found on these questions (Mans et al., 2015). These software uses algorithms to create models out of data which helps in analyzing the data. A few commercial applications are available, like Disco[2] and Perceptive Process Mining[3]. Since there exists a gap between the theory and the practice of Process Mining, the innovation, development and use of this technique in practice needs further research. This can be seen in hitherto performed researches in Process Mining, where the focus lays on the technique and algorithms itself (Caron et al., 2014). On the other hand, Mans et al. mentioned that Process Mining "*is mature enough to be applied to care processes of any type and of any complexity"* (Mans et al., 2015). Process Mining is a great tool to use for process discovery, conformance checking, prediction, and bottleneck analysis (Mans et al., 2015).

[2] http://fluxicon.com/disco/

[3] http://www.perceptivesoftware.nl/products/perceptive-process/process-mining

5 CARE MINING METHOD

This section will elaborate on the steps that are used to create the Care Mining method. The steps are a combination of both theories, VBHC and Process Mining. From VBHC the cost aspect will be used in order to generate more awareness about the costs (in terms of times) of processes (Mans et al., 2015). Process Mining will use event logs for the discovery of processes and can be accompanied with time averages from practice.

According to Kaplan and Porter (2011) TD-ABC describes a number of principal increments which should be taken for defining the cost of a cycle of care. These steps are described in Chapter 5 and indicate a criterion for determining a starting point till the ultimate determination of costs of each process in a care cycle. Herewith, process models will be created for an overview of existing performed activities. To create process models with the right and real data Process Mining will be used in these steps. Bozkaya et al. (2009) argued how analyses can be made from system's data with Process Mining (see Chapter 4.3).

5.1 CARE MINING METHOD STEPS

For the Care Mining method, not all the described steps are used to combine the methods of VBHC and Process Mining. The next section explains why and how these steps are included and performed. In Table 1 an overview can be found on these defined steps.

1. Select a medical condition

According to Kaplan and Porter (2011) the first step in creating quality improvements and cost control is by defining a scope. Another paper by James (1989) discusses to begin with finding a process in order to create measurements and analyses from. For the health care sector this will be done by selecting a specific medical condition or patient population. With this scope the researcher makes sure that costs measurements per patient can be created more effective and efficient (Kaplan & Porter, 2011). In order to create the appropriate measurements, the researchers need to have information about the condition and its performed processes. This information can be obtained by conducting interviews with the managers of the concerning departments and the medical specialist(s) of the chosen condition. The next section sub-steps will elaborate on the needed actions to be taken after the selection of a medical condition.

1.1 Gain knowledge about the medical condition

In the TD-ABC method, this step was not included. Before defining a scope one first needs to know more about the chosen medical condition. Therefore, the first sub-step is defined with the help of experts. This sub-step prepares the researcher to gain information and knowledge about the processes and activities of the selected medical condition. It is an important step to build a knowledge base for the next steps and to make appropriate analyses. With this step the researcher can understand and interpret the processes and activities conducted in a particular care cycle of a medical condition. To gain knowledge interviews are held which is described below.

Prepare interview

Every care cycle of a medical condition has a different start and end activity. This is the case due to differences in patients and complications which could occur. To define where a care cycle starts and ends interviews are held to ask the manager(s), of the particular department, and/or the medical specialists, of the condition, about the conducted activities. The managers are the people with knowledge about the processes and can help in creating process-maps. The medical specialists (i.e. nurses, physicians, etc.) are the group of people who have direct contact with patients and can give more in-depth information about the practical performances of the processes.

The interview questions are formed with previous received information (i.e. a build knowledge base with process-maps, documentations about the condition, etc.)

Conduct interview

As mentioned before, the interview is held with managers and or medical specialists. It is necessary to talk with people who have direct contact with patients to translate the predefined processes with the real situation. This group can also already state where they find bottlenecks or improvements in the current working method, which can be taken into account when analyzing the dataset.

After or during the interview the researcher should be able to determine the scope of the care cycle. When defining a start and end activity predefined process models or process definitions, achieved in previous sub-step, can be used in the interviews with the interviewees.

When the start and end activities are defined, it should be communicated with the concerned people (i.e. the interviewees). The approval is done when both parties, the researcher and interviewees, agree on what should be defined as the main start and end activity.

1.2 Define the scope of the research

This step defines and describes the scope and the purpose of what should be analyzed. The latter is important to define a specific scope and will help in stimulating medical specialists in reflecting on the performance of specific patient-related activities (McLaughlin et al., 2014). With the defined medical condition the final scope will be defined with demands from the medical specialists. These demands consists of questions which should be answered when analyzing the process models (Prodel et al., 2015). Without these questions, it is not clear what process models should be created and what should be the focus when analyzing the models. These questions could come from the medical specialists themselves or can be prepared together with the researcher.

After approval of the scope the performed activities in this scope should be made clear, which is described in the next step.

2. Define Care Delivery Value Chain (CDVC)

The second step is aimed at specifying the care delivery value chain (CDVC) and is also described in the TD-ABC method as second step (see Chapter 3.3). The CDVC make sure that the principal activities, which are conducted when delivering care to patients with a certain medical condition, are defined (Porter & Teisberg, 2006). In the CDVC the locations and involved medical specialists can be included, to create a complete overview about where and who performs certain activities. The CDVC focuses on the entire care cycle, rather than just individual processes (Porter, 2008). It also helps to identify relevant measurements, which is used to measure outcomes from.

To create a CDVC the next steps are defined with experts.

2.1 Gain information about care cycle activities

To create process models one should know what the performed activities are. This is possible by taking interviews or by direct observations on the concerned department(s) (Mans et al., 2015). This is also necessary for a better understanding of the dataset which will be acquired in the next step. The interview is held with the same group of people as mentioned in step 1. This interview can be taken with sub-step 1.1 Gain knowledge about the medical condition, which will safe time to make other appointments for another interview.

For this interview or direct observations, the CDVC of Porter and Teisberg (2006) will be used as a format. In Chapter 3.3 an explanation about the CDVC is given along with a depiction of it (see Figure 1123). The value chain includes all activities in the process of care that potentially brings value to the patient, from prevention to managing care (i.e. treatment and aftercare). The value chain connects patient needs to outcome indicators by defining the activities that are crucial for achieving most of the health benefits from. Figure 11 showed that the delivery chain consists of four layers. The first three layers, which are *Informing & Engaging*, *Measuring*, and *Accessing*, define the medical aspects around care for the patient. The last layer, which is *Care Delivery*, defines the activities that are stored in a EHR-system and contains measurable activities (Porter & Teisberg, 2006). The care delivery value chain for any medical condition starts with prevention and screening and ends with ongoing management of the disease. Health care providers do not always know beforehand what the next patient will need and therefore delivery is organized around general purpose units that can handle any problem in their discipline (Porter, 2008), which is defined in the fourth layer of the CDVC. Therefore, the last layer is used in this method to define the primary activities performed in the scope and can be seen in Figure 17.

Care Delivery (What activities are performed at each stage?)						
Typical path of patient care	Monitoring/ Preventing	Diagnosing	Preparing	Intervening	Recovering/ Rehabing	Monitoring/ Managing

Figure 17: The Care Delivery layer from the Care Delivery Value Chain (Porter & Teisberg, 2006)

2.2 Define performed activities

After the interview or direct observations, the main activities can be defined and filled in the CDVC template. Afterwards, this step will be used as a reference model to map the defined activities with the created process models for further analyses. This is done after step 3.

<u>3. Develop process-maps</u>

This third step describes how to model a process model with the use of the defined CDVC. The process models include the paths that patients follow, as they go through a care cycle. The following steps are defined with experts to create process-maps.

3.1 Identify and define dataset

Before creating process models a dataset with the relevant data is required, in order to deal with even logs and or specific patient groups of processes (van der Aalst et al., 2012). This sub-step and the next sub-step are additional steps added to the Process Mining Process of Bozkaya et al (2009). The reason is that in the Process Mining Process the researchers assume that a dataset is already available. In the case of this research the description of the acquisition of a dataset is required to describe a complete method. Also, often people extract data without having questions beforehand and data extraction is often done by people not analyzing the data. Data should be extracted through "trial and error" instead of a demand-driven approach (Mans et al., 2015). Therefore, this sub-step first begins by elaborating on identifying and defining the dataset. Figure 18 depicts an extension of the fourth layer of the CDVC, which describes where the primary activities should be obtained from. This extension has been made because Porter and Teisberg (2006) did not elaborate on the questions where information should be retrieved from, but only on the questions what should be retrieved in their template.

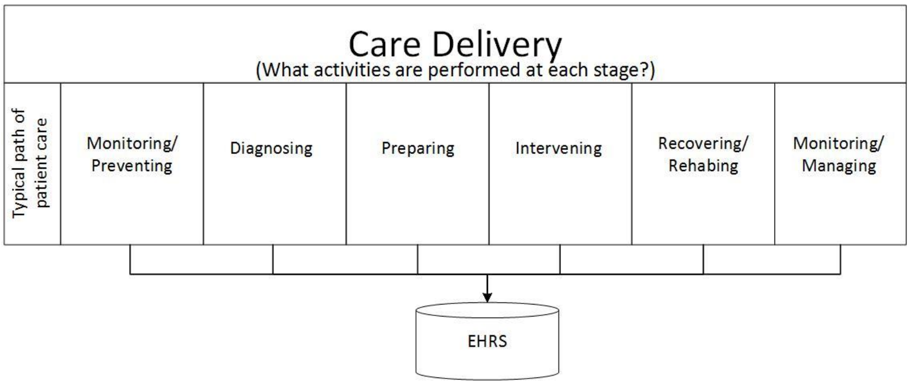


Figure 18: The extended template of the activities from the Care Delivery layer stored in an Electronic Health Record System (EHRS)

The first step to carry out is to identify data elements from databases, where potential patient-related data could be stored in. This means that relevant data should be identified or recognized in which database it is stored in. This is necessary because in a care cycle data could be stored in different departments and therefore also in different EHR-systems (Mans et al., 2008). A depiction of how care delivery is stored in different databases can be seen in Figure 19.

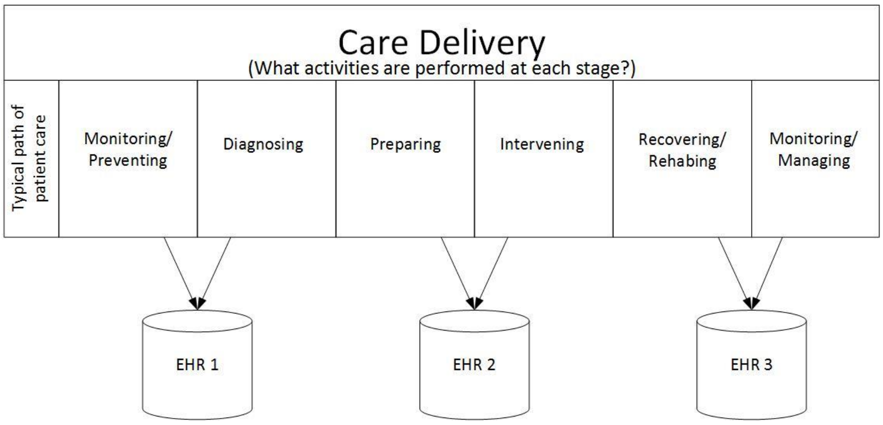


Figure 19: The activities from the Care Delivery layer in different databases

Note that not all data is measurable and therefore not necessary to retrieve. This has been determined after discussing this step with experts. Measurable data is performed in a period of time and has a start and end time (van der Aalst et al., 2005). For example, measuring the aftercare of an Obese patient. After this patient group is treated, they become an outpatient. This means that they will be treated by other health care organizations. The aftercare, when patients are a combination of in- and outpatient, could also last up to five years (depending on a specific case). In this case there is no data available. Another example of aftercare is when, for a certain condition, complications occur. Complications are a factor which could not be determined beforehand. Therefore, each medical condition must be examined whether a certain activity in the fourth layer of the CDVC could be measured or not.

Therefore, measurable data should be defined. To define which data is needed an interview needs to be held. This interview should be held with persons who have access and/or are able to retrieve the necessary data from the EHR-system. In this case, it is more distinct to communicate which data (linked to an activity) is required. Data about the main activities, specified in step 2.1, and data from the measurable processes, written in the fourth layer of the CDVC, should be retrieved and is explained in the next sub-step.

3.2 Obtain patient-related data

This sub-step describes the acquisition of the dataset with patient-related data. The data that will be obtained should be related to the questions defined in step 1.2. It needs to be obtained by someone who has access to and is able to retrieve data from the EHR-system (e.g. data-analyst) (Mans et al., 2008). Data can be delivered in Excel files and by other means. Figure 21 shows the next step in the extended template, which depicts that the activities stored and recognized in the database(s) are retrieved in the form of a dataset.

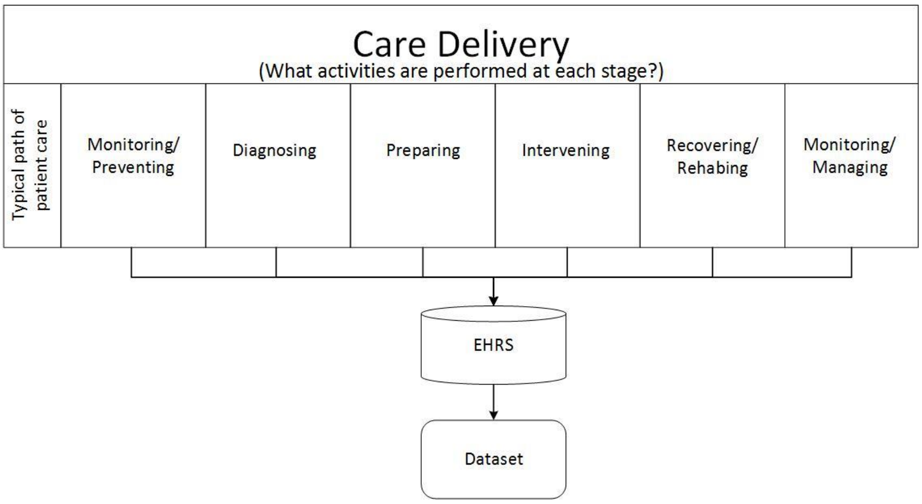


Figure 20: The extended template of the activities retrieved from an EHRS in a dataset

In the case when the activities are stored in different databases, the data is then retrieved in different datasets. A depiction of how activities, stored in different databases, are retrieved in different dataset can be seen in Figure 21. These datasets will need to be linked together to import and analyze the data in a Process Mining tool and is described in the next step.

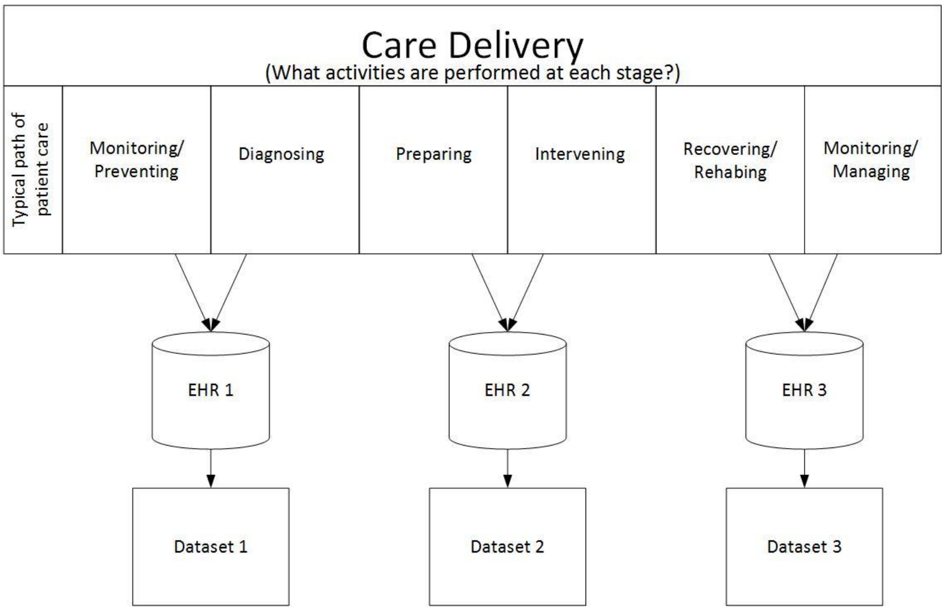


Figure 21: The activities retrieved from different databases in different dataset

3.3 Log preparation

From this sub-step the Process Mining Process will be used to create the process models. The dataset should eventually be imported into a Process Mining tool and therefore, this sub-step prepares the dataset to be able to be imported into the tool. The first step is linking datasets, in the case when different datasets were retrieved. In this way events from the different datasets are linked to each other in order to create one pathway in one dataset (Caron et al., 2014) (see Figure 22).

The second step is clustering the activities in the linked dataset. This step is optional for small datasets, but required in large datasets. Clustering activities improves the quality of the analyses (Rebuge &

Ferreira, 2012). It avoids spaghetti-like processes and makes sure the created process models are readable. The clustering process of Rebuge & Ferreira (2012) is used.

The third step in preparing the dataset is to make sure that event logs can be extracted from the dataset. In order to do so, three aspects should be present in the dataset, namely: case identifications, activity names, and timestamps (van der Aalst et al., 2012). Case identification (case id) is a unique numeric code which is linked to one case. An example of a case id could be a patient number which is specifically linked to one patient. Process Mining uses this id to create process models per case (i.e. a patient) and to calculate certain aspects, such as: patient numbers, number of activities per patient, etc. Activities are linked to case ids and therefore the names should be equivalent to each other. Meaning, the activities which are the same should have identical names defined. Timestamps are linked to a patient number and performed activity. It preferable includes dates of when an activity was performed (start and end date) and the time when an activity was performed (start and end time). When creating process models, these timestamps are used to create paths between activities and to calculate performance times of activities.

The fourth step is transforming the (clustered) dataset into the required extension (Caron et al., 2014; Bhattacherjee, 2012). The extracted dataset, from an EHR-system, could have different extensions. To import the dataset into a Process Mining tool it is important to consider which extension is required. If needed, the dataset needs to be transformed. This could be in any format, for example: Comma Separated Values (CSV) file, a MXML or XES format, etc.

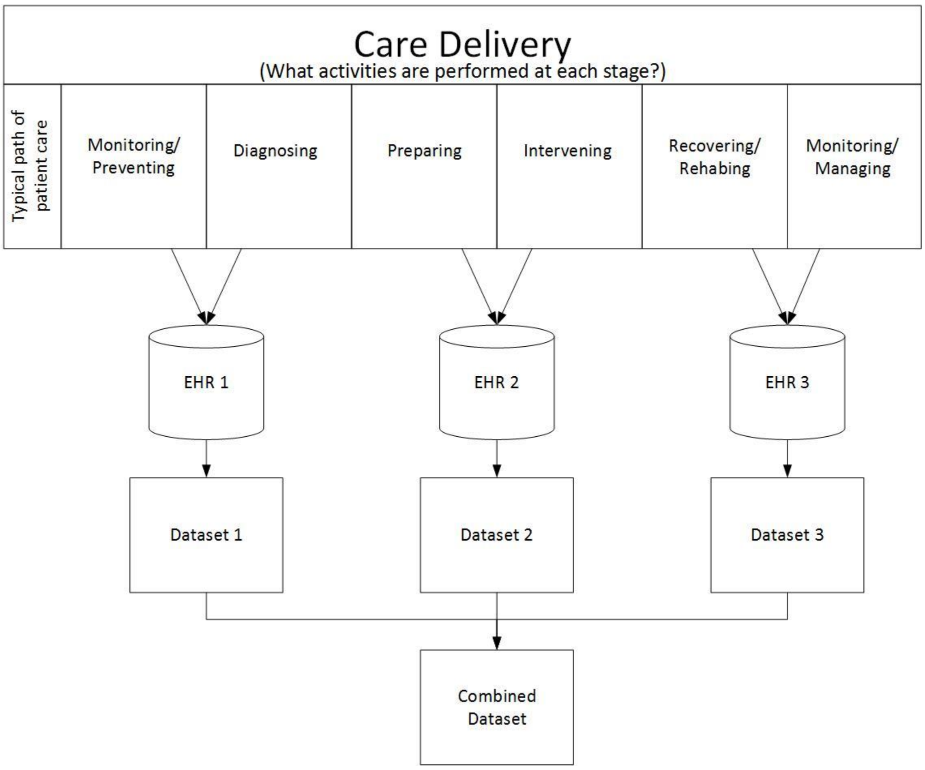


Figure 22: Event logs from different datasets linked into one dataset

3.4 Log inspection

In this sub-step an inspection of what the dataset means and contains is done. It makes sure that the created models are a true representation of the events (Caron et al., 2014). Also, in this way data tampering must be prevented and ensures to record the events timely and precisely (Mans et al., 2015). This step is in preparation for analyzing the process models with the appropriate statistics (i.e. activity performance times, activity paths, activity numbers, etc.) (Bozkaya et al., 2009). The focus should be on the existence of ids, activity names, and timestamps in the data per case. Also, it is important to check whether the data is in the required scope (i.e. in the agreed period). If not, the case(s) outside the scope should be removed. It is also important to notice if there are empty cells or inequivalent activity names.

When performing this step the researcher already gets to know the dataset more in-depth, which helps in understanding the dataset better and eventually for analyzing the process-models (Bozkaya et al., 2009).

3.5 Create process model

At this point the dataset should be ready to create process models from and will be imported into a Process Mining tool. A tool will make sure that information from the created models can be analyzed in a structural manner and eventually will save time and insures the accuracy of creating a model (Caron et al., 2014).

The data should already be prepared in the right extension in sub-step 3.3. After the import of the dataset in a Process Mining tool filters on the data must be used to define the type of model one wants to create. The filters ensure to clean up the data and to focus the analysis that must be performed (Rebuge & Ferreira, 2012). The process models will be created according to the defined questions in step 1.2.

4. Perform Analyses

This step analysis the created process models on correctness in accordance to predefined process definitions and prepared questions (from sub-step 1.2). The process-models are analyzed for any bottlenecks and the time performed for each process is estimated. This step is comparable with the Performance analysis step of Bozkaya et al. (2009) and the sixth step of the TD-ABC steps of Kaplan and Porter (2011) (i.e. Estimate the practical capacity of each resource provider). The fifth step (i.e. Estimate the cost of supplying each patient care resource) of the TD-ABC method is not included in this method, because this research does not focus in the financial part of the processes.

4.1 Process model analyses

After a model is created the researcher together with data-analysts and project leaders can analyze the information from the created model. This model can be analyzed on the efficiency, quality and conformance. Analyzing efficiency, generally, includes identifying bottlenecks, the number and duration of activities, and performance comparisons between activities with the same diagnosis. Quality and conformance analyses could include evaluations of medical rules, model-based conformance analysis of conflicting events and their causes (Caron et al., 2014).

5. Discussion and Validation

This is the final step in the Care Mining method and reports and discusses the results of the made analyses. Both the TD-ABC method of Kaplan and Porter (Kaplan & Porter, 2011) and the Process

Mining Process method of Bozkaya et al. (2009) did not elaborate on reporting the results and is therefore added to this model with the help of experts.

5.1 Report the analyses

The findings from the analysis are presented in a report with the process diagrams from the used Process Mining tool. In this way, the process owner obtains a clear picture of the reality of the process. The results are described with a PDD (focusing on the deliverables). In this report the Plan Do Act Check (PDAC) steps are described. In the Plan phase the predefined questions and or objectives are elaborated on. This phase ensures and checks that the analyses were made with the right questions and or objectives. In the Do phase the obtained measures are reported next to the questions and or objectives. The Check phase compares the found measures with the defined CDVC from step 2, to recognize differences and maybe bottlenecks. Finally, in the last phase Act the needed modifications or improvements are described or discussed.

5.2 Transferring the results

Finally, the reported analyzed need to be discussed with the concerned people. This step is described by Bozkaya et al. (2009) and by Mans, Schonenberg, Song and van der Aalst (2008). The concerned people are the ones who are involved in the process.

No.	Theory	Sub-step	Method/techniques/metrics	Argumentation	Reference
Main step					
1.	TDABC	**Select a medical condition**			Kaplan & Porter, 2008
1.1		Gain knowledge about medical condition	Interview: with physicians, nurses and managers (or project leaders)	This step is needed to know what the chosen medical condition is about and what the main activities are. This will help in defining a start and end activity.	Expert opinion
1.2		Define the scope of the research	Interview: with physicians, nurses and managers (or project leaders)	This step is the final scope of the research and is performed to make sure that analyses are made according to certain demands.	McLaughlin et al., 2014; Prodel et al., 2015
2.	TDABC	**Define Care Delivery Value Chain**			Kaplan & Porter, 2008; Porter, 2008
2.1		Gain information about care	interview/ direct observation	In this step information needs to be gained about the primary	Expert opinion; Mans et al., 2015

		cycle activities		performed activities. This increases the definition of the performed care cycle for the researcher and all the involved people.	
2.2		Define performed activities	Interview	During or after the interview/direct observations the main activities can be defined and filled in the CDVC template.	Expert opinion
3.	TDABC	**Develop process-maps**			Kaplan & Porter, 2008
		Identify and define dataset	Interview	Before obtaining data, the researcher should know in which databases (identify) patient-related data are stored in and what kind of data (define) is needed.	Van der Aalst et al., 2012; van der Aalst et al., 2005
3.1		Obtain patient related data	Interview	The acquisition of the dataset with patient-related data.	Expert opinion; Mans et al., 2008
3.2	PM	Log preparation	Interview	This step is performed to make sure that the obtained dataset can be imported in a Process Mining tool.	Expert opinion; Bozkaya, 2009; Caron et al., 2014; Rebuge & Ferreira, 2012; Van der Aalst et al., 2012
3.3	PM	Log inspection		Inspection of what the datasets means and contains is performed by the researcher to improve the quality of the analyses in step 4.	Bozkaya, 2009; Caron et al., 2014
3.4	PM	Create process model		The activity model is created by using different filters.	Expert opinion; Caron et al., 2014;

					Rebuge & Ferreira, 2012
4.	TDABC	**Perform Analyses**			Kaplan & Porter, 2008; Bozkaya et al., 2009
4.1		Process model analyses		This step analyses the process models for any bottlenecks and estimates the time performed for each process, with a focus on the defined questions.	Bozkaya et al., 2009; Caron et al., 2014
5.	TDABC	**Discussion and Validation**			Expert opinion
5.1	TDABC	Report the analyses		The results of these steps are reports. The results can be described with a PDD (focusing on deliverables)	Expert opinion
		Transferring the results	Meeting	After reporting on the analyses the results should be discussed with the involved people of this research, to reflect and discuss the CDVC defined in step 2 and the created process model of step 3.	Bozkaya et al., 2009; Mans et al., 2008

Table 1: An overview of the Care Mining method

5.2 GENERAL EVALUATION OF THE METHOD

This paragraph will elaborate on the validation of the method. The validation is the internal validation of this research. Because of this validation, improvement points of the method will be discussed in this paragraph.

The validation on the method has been performed with two experts, namely: a medical informatician and a process miner. These experts were involved to identify weaknesses and to improve it. These improvements were discussed in an iterative walk through sessions, to introduce the proper improvements. Walk through sessions were chosen, because it can be used in the development process and before the usability testing (Jeffries & Desurviret, 1992). The following points were discussed and proposed as improvement points.

The method consists of 5 main steps and each step is divided into sub-steps. These sub-steps describe in detail how the particular step should be executed. In Table 22 an overview is created of the previous Care Mining method with its improvements.

Previous method steps	**Improvements**	**Reasoning**
1. Select a medical condition	No improvements	/
1.1 Gain knowledge about the medical condition	No improvements	/
1.2 Define start and end process of medical condition	**1.2** Define the scope of the research	An additional and more refined scope is introduced, which make use of questions from the hospital. This method makes use of questions to create process models.
2. Define Care Delivery Value Chain (CDVC)	No improvements	/
2.1 Gain information about care cycle activities	No improvements	/
2.2 Define performed activities	No improvements	/
3. Develop process-maps of each activity	**3. Develop process-maps**	The models will be created for each process performed in the care cycle, instead of for each activity. Therefore, this step is changed in to Develop process-maps, which indicates that a map of process models will be created.
3.1 Obtain patient-related data	**3.1** Identify and define dataset	At first the dataset, where the required patient-related data is stored in, should be identified. Afterwards the necessary data from the dataset should be defined.
3.2 Log preparation	**3.2** Obtain patient-related data	This step is the same, but will be performed secondly.
3.3 Log inspection	**3.3** Log preparation	This step is the same, but will be performed thirdly.
3.4 Create activity model	**3.4** Log inspection	This step is the same, but will be performed fourthly.

3.5 Control flow analysis	**3.5** Create process model	For this step, only the name is changed from ***Create activity model*** to ***Create process model***. This step is the same, but will be performed fifthly. The Control flow analysis is performed in step 4.
4. Obtain required estimates	**4. Perform Analyses**	Instead of directly obtaining estimates of the created process model, the model first needs to be analyzed.
4.1 Estimates for each process step	**4.1** Process model analyses	Due to the change in step 1.2 the process models are created according to questions. In this case the analyses will not be created for every process step, but just for the created process models. Moreover, this step does not make estimates, but takes the real data from the process models.
4.2 Estimates for the cost of supplying each patient care resource	Sub-step is removed.	This sub-step has been removed because this method does not focus on the costs of performing care. This could be done in a future research.
5. Compute total costs over each patient's care cycle	**5.** Discussion and Validation	Due to the removal of step 4.2 the total costs, as described by Kaplan & Porter (2011), cannot be calculated. Instead, the made analyses from step 4.1 will be discussed and validated with a medical specialist(s).
5.1 Calculate resource capacity costs	**5.1** Report the analyses	Calculating the capacity costs of the resources is also a step which includes using costs. Therefore, this step is replaced by reporting the analyses. This should be done to explain and transfer the results.
5.2 Calculate total cost per patient care cycle	**5.2** Transferring the results	For the last step, calculating the total costs is replaced by Transferring the results. This step is the last phase described by Bozkaya et al. (2009) and transfers and discusses the results with the concerned hospital personnel.

Table 2: Evaluation of the Care Mining method

In total thirteen main- and sub-steps, from a total of eighteen, were included, changed or removed after validating the method with the two experts. This means that one sub-step was included (e.g. 3.1), eleven main- and sub-steps were changed (e.g. 1.2, 3, 3.2, 3.2, 3.4, 3.5, 4, 4.1, 5, 5.1, 5.2), and two sub-steps were removed (e.g. 3.5 and 4.2).

When looking at the new order of the method (see Figure 23), it should be more clear, efficient and easy in its execution. The method begins with defining a scope by selecting a medical condition (1. Select a medical condition). The medical condition is often already given by the hospital who want to review and analyze their performed processes. For the researcher to analyze the medical condition, knowledge about it needs to be created with available information (e.g. step 1.1). Then a more refined scope is created with specific questions the hospital has for creating the overviews (e.g. step 1.2). The whole first step is necessary for creating a base for the research and will be the guidance when performing the next steps. Afterwards, step 2 has not changed and is needed for defining the performed activities. This is then used in the next step for identifying and defining a dataset (e.g. step 3.1) and for analyzing the created process models (e.g. step 4). Therefore, this step can be considered as the golden standard. Step 3 is the core of this method and determines the direction of the application and outcome of the used case. After the creation of the process models, according to the predefined questions, analyses are performed to answer these questions (e.g. step 4). Finally, in step 5 the results are reported and discussed with the concerned hospital personnel. In this way, this method should be able to create an overview or overviews of hospital's processes.

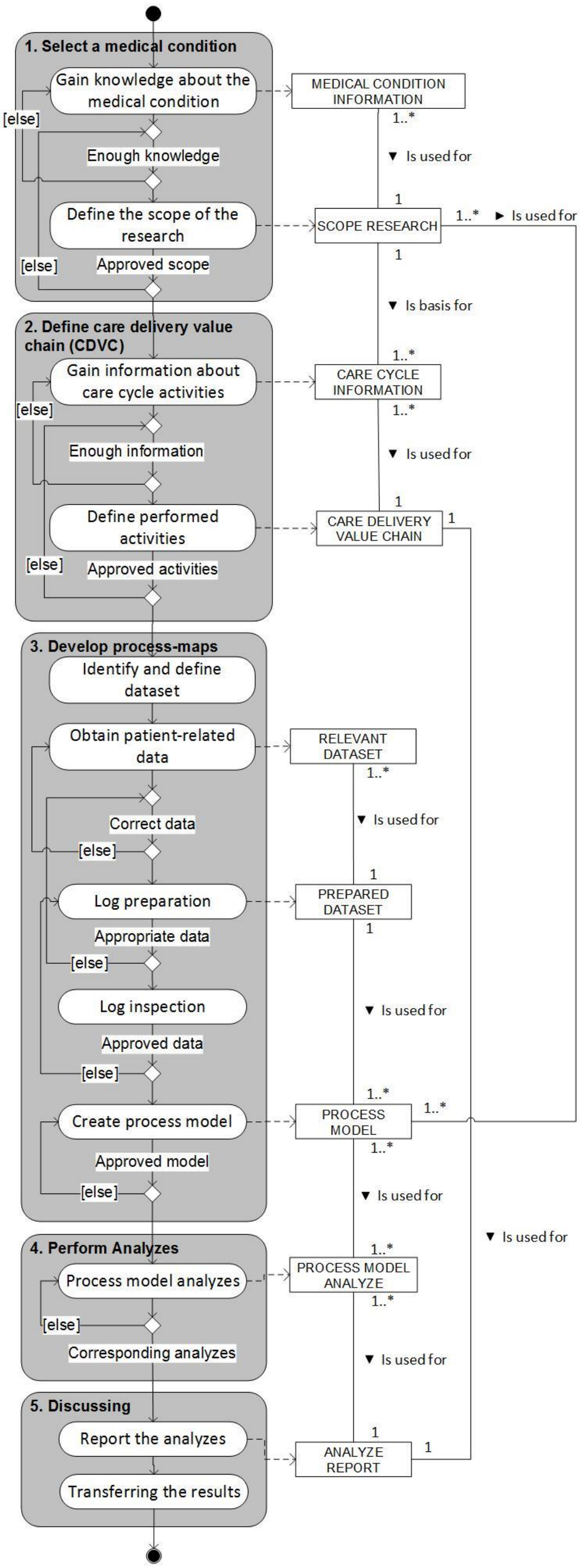


Figure 23: The Care Mining method after validation

6 APPLICATION OF THE METHOD: CASE STUDY

This chapter will describe the results of the application of our Care Mining method to two cases namely Breast Cancer & Colorectal. The cases will be described first and after that each step in this method will be described for each case.

Meetings and discussions have been conducted at different hospitals to accommodate the Care Mining method. However, not all hospitals were ready to work with this research. This had several reasons, among which: 1) that in some cases hospitals should still develop more awareness about the added value of the integration of the combination of Process Mining and VBHC for their organization, 2) in many cases hospitals were excited about this research, but in a short space of time there was no room for it. In the end, two cases were acquired and used, where there was a space and time for this research at the time when this research started.

6.1 CASE: BREAST CANCER

The first case study took place in a health care organization in the South of the Netherlands. This hospital is a general hospital, which means that it has a concentration of facilities for regular patient care for research, treatment and care. In addition, in a general hospital, both doctors and nurses are trained (volksgezondheidenzorg.info, 2015). In this paragraph the hospital will be referred to as Amhos. Amhos is hospital with more than 4000 employees (Goedee, Suttorp, & Lieshout, 2014) and has the most outpatient visits from all the hospitals in The Netherlands (Ministerie van Volksgezondheid, 2016; Smit, 2016). The clinical pathway within this case study is about breast cancer and is called: Breast Cancer.

Breast cancer usually arises in the mammary glands of either of the two breasts. The number of new cases of mammary carcinoma increases (Integral Cancer Center in the Netherlands). In 2002 it was around 11,500, with more than 14,000 people in the Netherlands diagnosed with mammary carcinoma. This form of cancer is the most common form among Dutch women (Marianne Heins et al., 2015).

At the Breast Clinic of Amhos, a team of physicians, nurse practitioners breast care and breast care nurses. is working. The medical team often consists of a surgeon, radiologist, pathologist, radiation oncologist, medical oncologist and nuclear physician.

At the request of the hospital, it has been asked to create an overview of the pathway Breast Cancer with the basic numbers of patients and lead times, using Process Mining.

1. Select a medical condition

1.1 Gain knowledge about medical condition

To gain knowledge about breast cancer and the execution of the treatment, information was sought on the website of the hospital first. This was done to prepare for the first interview with the Business manager, who asked to create an overview of the pathway Breast Cancer. On their website, there is a description about the execution of the treatment process in general, which is from inspection until creating a treatment plan. Every patient is different and therefore needs a different treatment plan. This raises the question how an overview could be made for the Breast Cancer pathway. The pathway breast cancer has a high degree of variety in the operations order. A standardized process is substantially not perceptible. A large group of patients, therefore, passes through a unique path. With respect to the processing times, this has the consequence that at a more detailed level a high degree of variety in processing times occurs.

Also, a predefined or theoretical model was given by the manager and depicts the important activities (see Appendix 12.2).

1.2 Define the scope of the research

The process has a wide variety of entry points and almost all existing transactions also occur as a starting point in the pathway. To report on the turnaround times an aggregation of activities and a well-defined start and endpoint is created. As a "fixed" starting point for the turnaround times the operation “1st clinic visit” was chosen. This moment comes the closest to the start time of the theoretical model (see Appendix 12.2). For an end process, we discussed that two activities were most common, namely: Chemo and Surgery. In step 2, a depiction of the order of the main activities can be seen in Figure 25.

The questions the hospital had for analyzing their activities were about two main subjects, namely: 1) creating an overview of the elemental number of patients in their pathway from 2014, and 2) creating an overview of the turnaround times of the most followed paths which is depicted in Figure 25.

<u>2. Define Care Delivery Value Chain (CDVC)</u>

2.1. Gain information about care cycle activities

For this step an interview was held with a manager and a physician, to ask about and define the main important activities. The activities mentioned in the scope were used as a starting point in this interview.

We found that activities around this care cycle are generally executed in a specific way of order (see Figure 25). There are two ways in which a patient ends up in a hospital, namely: 1) if something suspicious is seen in the population survey, one will be referred for additional research to a hospital's Breast Clinic; 2) if the general practitioner finds something suspicious, he will point the patient through. The patients are sent to the hospital via a letter of referral from the general practitioner. In this context, the patients are in the first activity of this scope, namely the activity "First Clinical Visit". This activity means that there is a conversation between the patient and a specialist about an investigation into possible breast cancer. After this conversation imaging diagnostics and pathological examination is carried out to investigate whether a patient has cancer or not. With the results of the two investigations a multidisciplinary consultation is planned, to discuss the appropriate treatment. In general, a patient can be treated with a chemotherapy or a surgery.

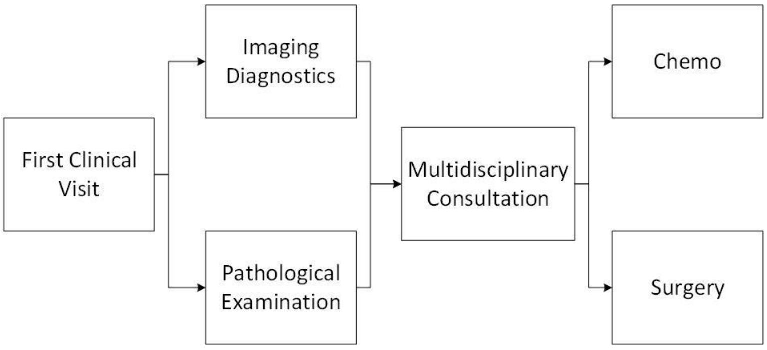


Figure 24: The scope of Breast Cancer Case

2.2. Define performed activities

In Figure 26 a depiction of the activities performed in the care delivery chain of Breast Cancer within the defined scope can be seen. Only the first four phases were filled in, because the activities chemo and surgery are part of *Intervening*. The activities Rehabing and Managing are ongoing processes and are difficult to measure in this research, due to a lack of time. Those activities could take a year or five to be able to measure its data.

Care Delivery (What activities are performed at each stage?)						
	Monitoring/ Preventing	**Diagnosing**	**Preparing**	**Intervening**	**Recovering/ Rehabing**	**Monitoring/ Managing**
Typical path of patient care	- Medical history - Control of risk factors (obesity, high fat diet) - Generic screening - Clinical exams - Monitoring for lumps	- Medical history - Determining the specific nature of the disease - Generic evaluation - Choosing a treatment plan	- Surgery preparation (anesthetic risk assessment, EKG) - Plastic or oncoplastic surgery evaluation	- Surgery (breast preservation or mastectomy, oncoplastic alternative) - Adjuvan therapies (hormonal medication, radiation, and/or chemotherapy)		

Figure 25: CDVC for Breast Cancer

The activities of the Care Delivery figure are all grouped under one denominator. Thus, the process models could be created more clearly and more accurately. For example, the activities under Intervening are among Chemo and Surgery.

<u>3. Develop process-maps</u>

3.1 Identify and define dataset

In this step, we took an interview with the manager and a business analyst, who has access to patient-related data. They described that data for patient numbers and turnaround times between first clinical visit and chemo and surgery are processed in the databases of their EHR-system and the Anatomical Pathology department. In the EHR- system all the activities from Monitoring/Preventing until

Monitoring/Managing are stored. In the Monitoring/Preventing part laboratory research is conducted by the Anatomical Pathology department, which uses a different database called PALGA. Therefore, this database was also used in this research. The data from both datasets cover a large part of the care cycle.

For this research, we needed the columns patient numbers, timestamps, and activity names from the main activities Monitoring/Preventing, Diagnosing, Preparing, and Intervening. We had defined that start and end times, or dates, from each activity should be present in the database, to create models with turnaround times.

3.2 Obtain patient-related data

From the EHR-system and PALGA data between 01-01-2014 till 31-12-2015, with the main activities, was extracted from their patientcare database in an Excel Binary File Format (.XLS)[4]. Due to difficulties to specifically extract only this data from the databases, we got two datasets with all the activities carried out in 2014 from the two EHR applications. In the Log Preparation step, the necessary columns were selected and placed in the right order in another file to keep an overview of the required columns. This process will be explained in the next section.

3.3 Log preparation

The dataset for the Process Mining tool was created in different iterations. First, the project started with the dataset from the EHR-system and later the database from PALGA was added, since the exact date and time were recorded in a certain process step. This was done to improve the quality of the results of the Process Mining project. The activities to create a formatted CSV was depicted in Figure 26 and these steps were performed four times by the Process Mining analyst.

The patient related dataset also contained activities from 2011, which was not relevant for the analysis of 2014. Therefore, the cases before 01-01-2014 were removed from the log file, to ensure that the start of relevant cases was preserved.

After the first log inspection, we noticed that there were 457 different activities in the dataset. Among these activities there were several activities that could be grouped (e.g. CT-Scan 'MRI Pelvis' and 'MRI Abdomen'), which essentially is a MRI scan. Grouping or clustering activities was also needed to generate the relevant throughput times. The method from Rebuge & Ferreira (2012) was applied to cluster the activities. This reduced the number of activities (e.g. events) from 475 to 27 (see Table 3). These results were validated with a specialist from Amhos.

Clustered operations from the EHR- system and PALGA datasets	
(Para)medical and support functions	Microbiology and Parasitology
Imaging Diagnostics	Not Diagnosis Treatment Combination Maintenance
Special artificial means and aids	Not Captured in Profile
Blood products	Surgical Operations
Chemotherapy	Other Laboratory Operations

[4] https://support.microsoft.com/nl-nl/kb/840817

CT-research	Other Therapeutic Operations
Day Care	Other Care Activities for Distraction
Diagnosis Treatment Combination Maintenance	Palliative Treatment
Diagnostic Activities	Pathology
Expensive Drugs	Clinical- and Emergency Room Visits
Hormone Therapy	Start Chemotherapy
Clinical Chemistry and Haematology	Nursing Day
Clinical Admission	Preparing Operation
Multidisciplinairy Consultation	

Table 3: Clustered operations from the EHR- system and PALGA datasets

Afterwards, a validation of the clustered activities was done on the simplification of the process steps with a nurse of Amhos. This method has already been applied and presented in the paper of Rebuge & Ferreira (2012). This step of simplifying the process steps is performed two times to create different abstraction levels.

3.4 Log inspection

The patient dataset had 5.854 patients and PALGA 1.743 patients. Also, we found 365 patient numbers who were in the PALGA dataset and not in the EHR- system dataset. This was strange because, normally, patients are referred to the Anatomical Pathology department from the patient database. A specialist from that department mentioned that there could be several reasons. The first one could be that the diagnosis treatment combination (which is a payment system for hospital care based on a standard treatment code) are closed in the EHR- system, because the patients are treated in a different department. In this case, it could be the case that those patients are not mentioned in the EHR- system. Another reason could be that patients are referred to Amhos by their primary care physician and have had a biopsy via the Radiology department. The results from the biopsy are processed in the EHR- system and the diagnosis treatment combination is closed. The problem in this case could be that the diagnosis treatment combination was closed before the results were sent. After the first inspection, a total of 6.219 unique patients were found in both datasets of patient related dataset and PALGA in 2014.

When both datasets (i.e. the EHR- system and PALGA) were linked, we could tackle this problem by creating another column with the origin of each patient. In this case, we created an overview with which the same patient numbers could be grouped.

3.5 Create process model

With the help of the Process Mining tool DISCO the care process is reconstructed. This results in the flow diagrams with corresponding numbers of patients and duration, which can be seen in step 3.5. The information generally provides direct new questions. With the aid of process mining software, the answers to the questions can be found then.

The linked dataset was imported into DISCO. The process models were created using filters in DISCO. With the filters the amount of activities could be reduced to maintain an overview. We limited the number of activities to seventeen. However, the pathway exhibits a high degree of variety (see Figure 26). Due to this case, it differs significantly from the theoretical model.

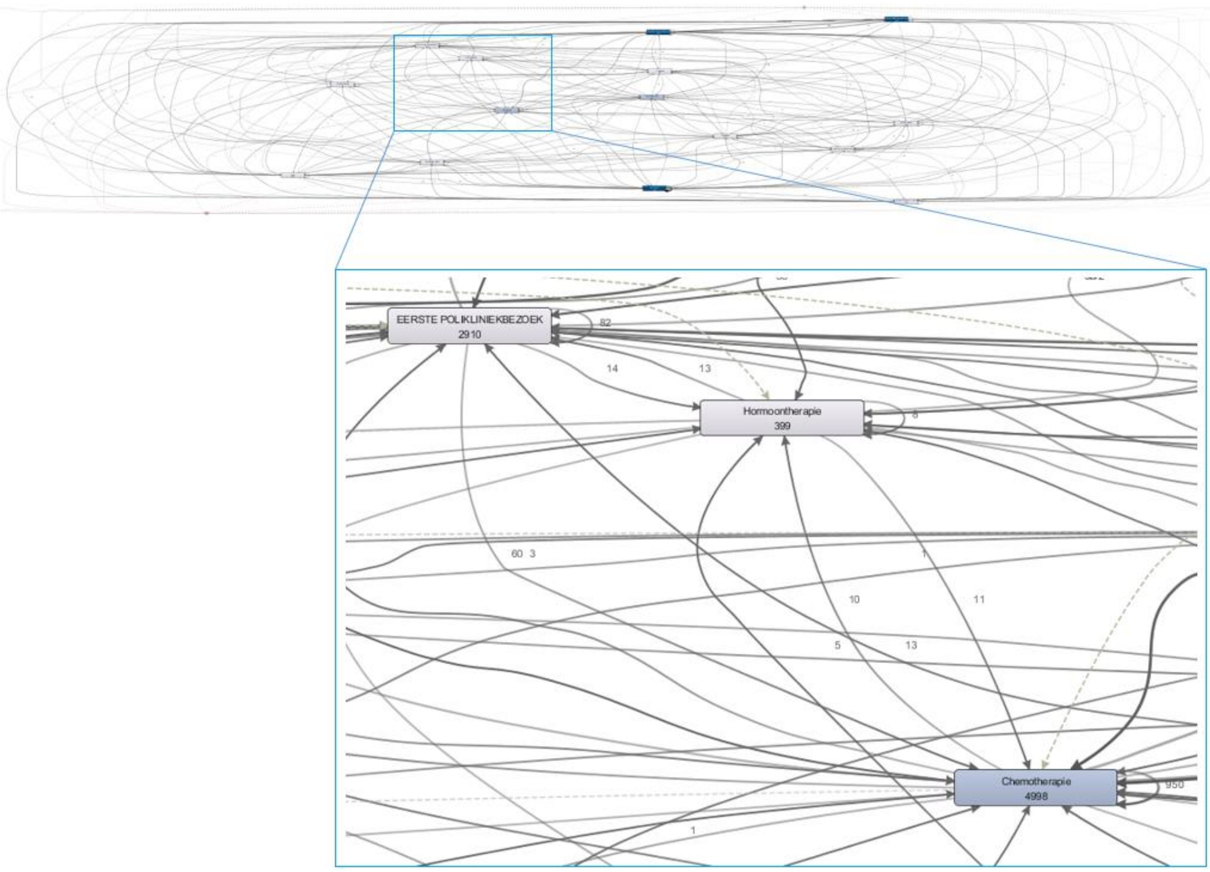


Figure 26: A zoomed-in depiction of the turnaround time between first clinical visit to Chemo

4. Perform Analyses

4.1 Process model analyses

The process models were created according to the two main questions of the hospital. The process of the clinical pathway of Breast Cancer was described in a multidisciplinary session. The medical specialists also supported the Process Mining analyst in the analyses by pointing out sequences in the process models. The following paragraph will describe the results of both analyzed questions.

Creating an overview of the number of patients who were treated in the year 2014:

This dataset is a combination of new patients and patients who were already treated in 2013. In the combined dataset, we found a total of 3.747 patients who were treated in 2014. We analyzed that 3.484 patients were stored in the EHR- system. Furthermore, four diagnosis codes were involved in this dataset, namely: diagnosis code 316, diagnosis code 317 (i.e. Benigne), diagnosis code 318 (i.e. Maligne), diagnosis code 811 (i.e. internal medicine). This code indicates the nature of the illness or disability in the form of a classification of diseases in the "Classification for Health & Safety and Social Insurance." Diagnosis code 316 was involved in ten patients, diagnosis code 317 was involved in 1.383 patients, diagnosis code 318 was involved in 1.735 patients, and diagnosis code 811 was involved in 736 patients. From all these patients, 380 patients had one or more diagnosis codes. All these number could be analyzed with the help of the dataset.

Creating overviews of turnaround times of the most followed paths:

For this question, we looked at the turnaround times within the predefined scope to include the important and most performed activities.

The first turnaround time we looked at was from the first clinical visit to imaging diagnostics.

In DISCO, we used a filter to select the two activities. The modeled process model is depicted in Figure 28 and Figure 29. In Figure 28 the process model is created with the number of patients involved in that process in one year. Figure 29Figure 28 shows the same process model with the turnaround times.

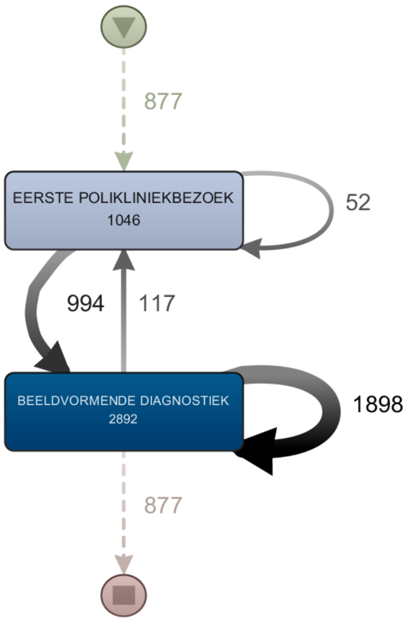


Figure 27: Number of patients between first clinical visit and imaging diagnostics

Figure 28 shows that 877 patients have a first clinical visit as a starting point and went through the imaging diagnostics activity, where the same 877 patients are going out from. Between these two activities more patients are involved, which is caused by the fact that these patients have a different starting activity. In one year 1046 patients went through a first clinical visit and 994 patients had imaging diagnostics as the next activity, after the activity first clinical visit. The thickness of the arrows depicts the number of patients that went through an activity.

Furthermore, 117 patients went from imaging diagnostics to first clinical visit for unknown reasons. A total of 52 and 1898 patients were in a loop in one activity had to be investigated further, because according to the predefined process model this should not be the case.

In Figure 29 the top number indicates the average processing time. The lower number indicates the number of patients which switches from one transaction transferred to the other. With this process, there are 6.4 days between "first clinical visit," and "imaging diagnostic", where 994 patients are involved in.

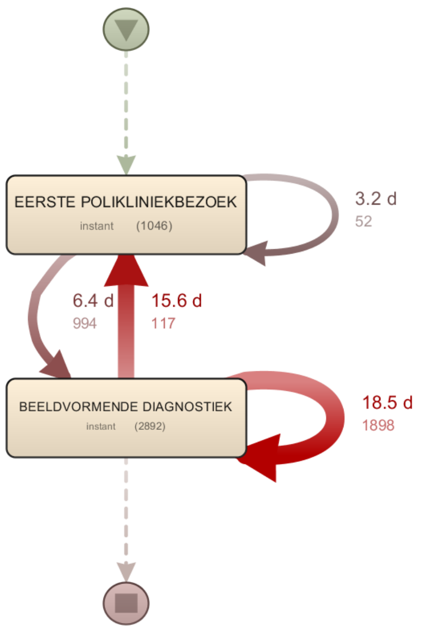


Figure 28: Turnaround times of patients between First Clinical Visit and Imaging Diagnostics

The next analyzed process was from First Clinical Visit to Pathological Examination.

In this process 396 patients went through (see Figure 30). The 517 patients in the first clinical visit could indicate that some patients from the year 2013 were still in this activity or it could be patients that went from the activity pathology back to the first clinical visit. This is also the case for the 3042 patients in pathology. A loop can also be seen in both activities, which means that 24 patients in first clinical visit and 2613 patient in pathology again go through the first clinical visit-activity. This is not possible, so this could be an error in the system.

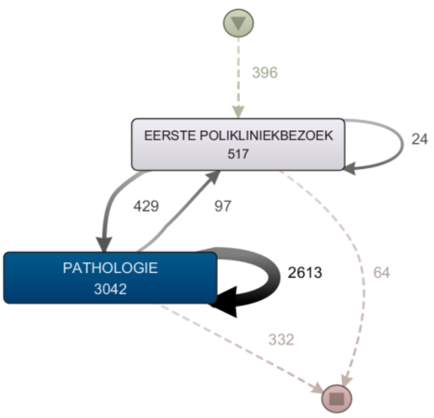


Figure 29: Number of patients between First Clinical Visit and Pathological Examination

Furthermore, it will take a patient approximately 6 days to go through this process.

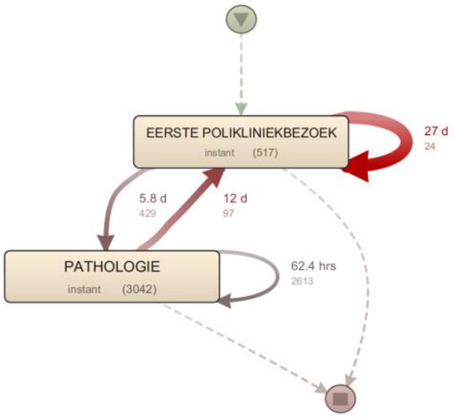


Figure 30: Turnaround times of patients between First Clinical Visit and Pathological Examination

The next process which was analyzed was from First Clinical Visit to Multidisciplinary Consultation.

It was analyzed that 381 patients first started with the first clinical visit and then went to the multidisciplinary consultation. On the other hand, a loop in both activities can be seen which indicates that patients did not go to the next activity, but went to the same activity again.

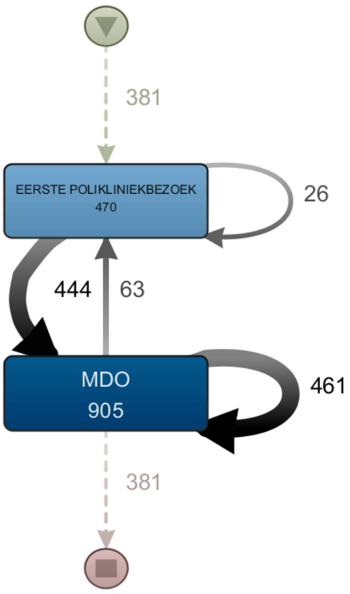


Figure 31: Number of patients between first clinical visit and Multidisciplinary Consultation

Also, around the 19 days was needed to get one patient through this process.

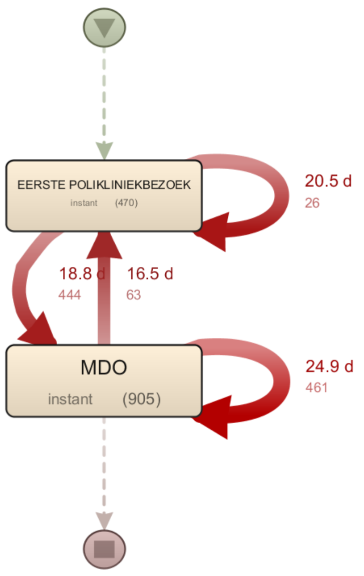


Figure 32: Turnaround times of patients between First Clinical Visit and Multidisciplinary Consultation

The next analyzed process was from First Clinical Visit to Chemo.

In total 198 patients straight went from the first clinical visit to a chemo. In this process, it also became visible that both activities had loops and that patients went back from a chemo to the first clinical visit.

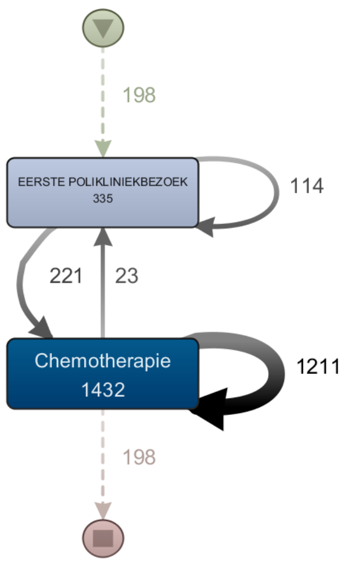


Figure 33: Number of patients between first clinical visit and Chemo

Furthermore, around the 18 days went through this process to treat one patient.

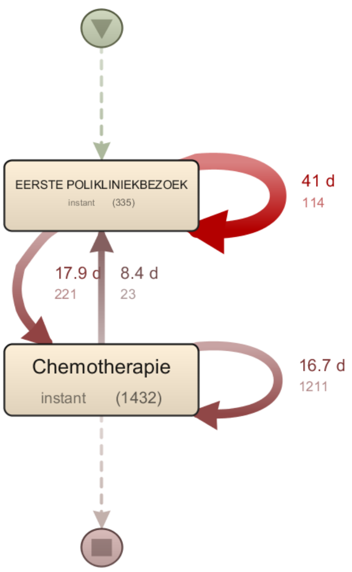


Figure 34: Turnaround times of patients between First Clinical Visit and Chemo

The next and last process is from First Clinical Visit to Surgery.

In this process, we analyzed that 285 patients underwent a first clinical visit and had a surgery afterwards. Both activities contained more patients than patients that go into the activity, which could indicate that those other patients have another start point than first clinical visit.

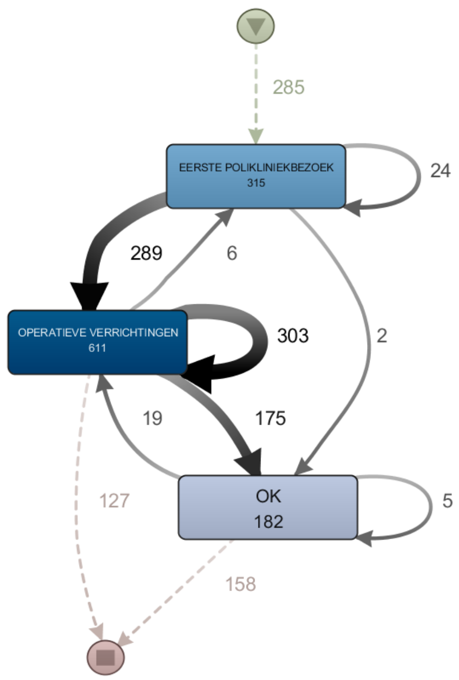


Figure 35: Number of patients between first clinical visit and Surgery

Lastly, one patient takes approximately 14 weeks to get a surgery from a first clinical visit.

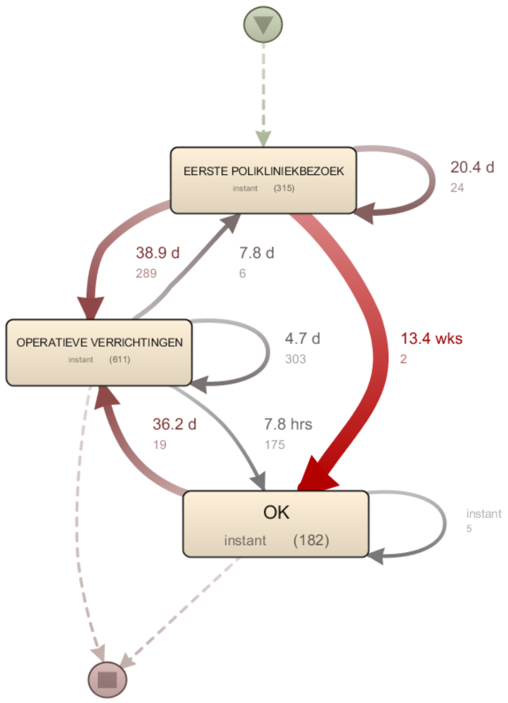


Figure 36: Turnaround times of patients between First Clinical Visit and Surgery

5. Discussion and Validation

5.1 Report the analyses

The created process models were discussed using a power point presentation. In this report, there are several snapshots recorded (see step 4). We discussed the information obtained from Process Mining with content experts. In this manner, the information from the analyses could be interpreted in a proper form.

The strength of using a Process Mining tool is that creating the process models are not time consuming (Corbey, 2008). The added value of Process Mining lies in the filtering, within the process, of the variants, which are not interesting for the clinical pathway. This improves the quality of the throughput times and makes is possible to zoom in on a patient level, instead of a Key Performance Indicator (KPI).

Everything could be measured. The only problem with this analysis was the comparison with the process model we received. In that model, other names for activities are used than in the EHR- system. That made it difficult to find out what activity, from the model, belongs to which activity from the EHR-system.

5.2 Transferring the results

The results of the throughput times of the clinical pathway were presented to the manager of the department. Then, the results were communicated to the doctors in a multidisciplinary meeting. The doctors reacted critical, since the results and analyses possibilities should be available within the current system (i.e. EHR- system). Afterwards, all the created process models and descriptions of it were sent to the manager.

6.2 CASE: COLORECTAL

The second case study took place in the same health care organization as the first case, namely: Amhos. In this case study the clinical pathway 'colorectal' was analyses.

At the request of the hospital, it has been asked to create an overview of the pathway colorectal with the basic numbers of patients and lead times, using Process Mining. Currently, Amhos did not have sufficient insight and control over the entire process.

Till now Amhos had no tool or clear method to get the right statistics. These statistics with lead times are used by DICA which is an eternal organization who uses the statistics to determine a pricing for the hospital. Amhos wanted to have more control over their own management information, without having to wait for the results of an external party. The results of colorectal for 2013 were surprisingly and therefore the hospital chose this medical condition to be analyzed for 2014.

<u>1. Select a medical condition</u>

1.1 Gain knowledge about medical condition

Firstly, knowledge for this case was gained from the hospital's website in order to understand the content and performance of the colorectal cycle. Secondly, a research report was studied, which was executed by a student and contained management information with turnaround times of 2013. This report was helpful to be used for mapping the current activities.

1.2 Define the scope of the research

During an interview, together with the department manager, it was determined what the first activity for a patient should be and the last activity. The last activity ended before any after-treatment, because after treatment can take a minimum of 5 years which is not included in a dataset of one year. For this case the hospital asked to analyze data of 2014.

It was determined that the starting point of the first activity is "First Doctor Visit". There was no theoretical model to validate the first activity, so this was done in the next step with other medical specialists. For an end process, we discussed that one is most common, namely: Surgery.

For this case the following two questions were analyzed, namely: 1) creating an overview of the elemental number of patients in their pathway from 2014, and 2) creating an overview of the turnaround times of the most followed paths.

The processes were analyzed on their turnover times. We also created the process models to see their progress between each other, which is described in the next steps.

<u>2. Define Care Delivery Value Chain (CDVC)</u>

2.1. Gain information about care cycle activities

To gain information about the activities in the care cycle of colorectal, interviews were held with the concerned internist oncologist, surgeon, stomach intestine liver specialist, and business analyst. These interviews were focused on the activities performed in the scope of the care cycle, as determined in the first step of this model. The analyzed processes are displayed in Figure 378. In total eight processes were analyzed.

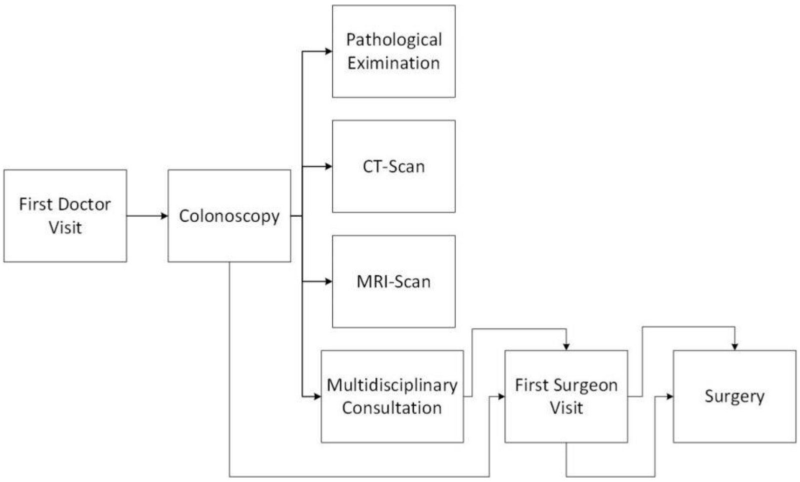


Figure 37: The scope of Colorectal Case

2.2. Define performed activities

The activities in this case study are the same as the activities of the case Breast Cancer (see Figure 26 and 39). The only difference is that the scope of this case study ends with activity Surgery, which means that the activity Chemo is left out. This was decided because the activity Chemo in this case study is relatively not a frequent performed activity.

Care Delivery
(What activities are performed at each stage?)

Typical path of patient care	Monitoring/ Preventing	Diagnosing	Preparing	Intervening	Recovering/ Rehabing	Monitoring/ Managing
	- Medical history - Control of risk factors (obesity, high fat diet) - Generic screening - Clinical exams - Monitoring for lumps	- Medical history - Determining the specific nature of the disease - Generic evaluation - Choosing a treatment plan	- Surgery preparation (anesthetic risk assessment, EKG) - Plastic or oncoplastic surgery evaluation	- Surgery (breast preservation or mastectomy, oncoplastic alternative)		

Figure 38: CDVC for Colorectal Cancer

3. Develop process-maps

3.1 Identify and define dataset

The activities in this case study are recorded in three databases, namely: PALGA, the EHR- system, and Dutch Surgical Colorectal Audit (DSCA) (DICA, 2016). PALGA is the data set of the pathological anatomical research. Within this dataset, it is determined whether a patient has cancer. This is recorded in every procedure, from resection to biopsy. the EHR- system is the medical catalog of Amhos, which should only contain patient data. DSCA is the database of the external registration/reporting, on which Amhos is settled. This is recorded by Integral Cancer Center Netherlands (IKNL) and three months after date, patient data is updated in their system.

3.2 Obtain patient-related data

The datasets with patient related data from PALGA, the EHR- system, and DSCA of the year 2014 were exported in an XLS-format. To analyze the datasets in the correct order in time, one dataset was created with all the datasets (e.g. PALGA, EHR-system, and the DSCA dataset) combined in it and ordered in time. To ensure that the dataset could be imported into a Process Mining application, the dataset was converted into a CSV dataset.

3.3 Log preparation

The creation of this dataset was realized with the performance of different processes. First, the dataset from the patient database and DSCA was used. The data from the EHR-system includes internal data from treated patients and the DSCA dataset includes external data, which is data registered outside Amhos, with results from primary colon cancer operations. In this way, the professional group gets insight into the quality of their own care and that of colleagues. The DSCA dataset was included in this research in order to report about the throughput times. Afterwards, the dataset from the Anatomical Pathology department was added because this is an important step where patients go through. Using this dataset, which is the Public Pathology Database (e.g. PALGA), we were able to analyze the process steps of the patients with its exact date and recorded time.

3.4 Log inspection

After the datasets were combined into one dataset, the first log inspection could be performed. All the data before 2014 was removed, to ensure that only data starting from 2014 until 2015 was inspected.

The combined dataset consisted of more than 30.000 records of data. With this amount, it would take days to generate models of the processes. Hence, process steps were grouped together for the simplification and creation of the required throughput times and process models with a Process Mining application. The cases and events were clustered together with the method from Rebuge & Ferreira (2012).

Hereafter, the clustered dataset was validated with nurses from the hospital. This was done to determine what processes could be clustered. The clustered dataset was validated again with Oncologist and business analyst of the department, also to check if the order of the process performance was in accordance to the reality.

3.5 Create process model

For creating process models DISCO was used for the simplicity of the application. Filters were used to create the models. The filters were used by selecting the concerning process steps and per throughput time this is being adjusted to the concerning activity. For example, to create a process model from “First Doctor Visit” to “Colonoscopy”, these two activities were selected (see Figure 40). The Process Mining Application then created the model with the throughput times in days and hours between the two activities, which can be seen in step 4.1.

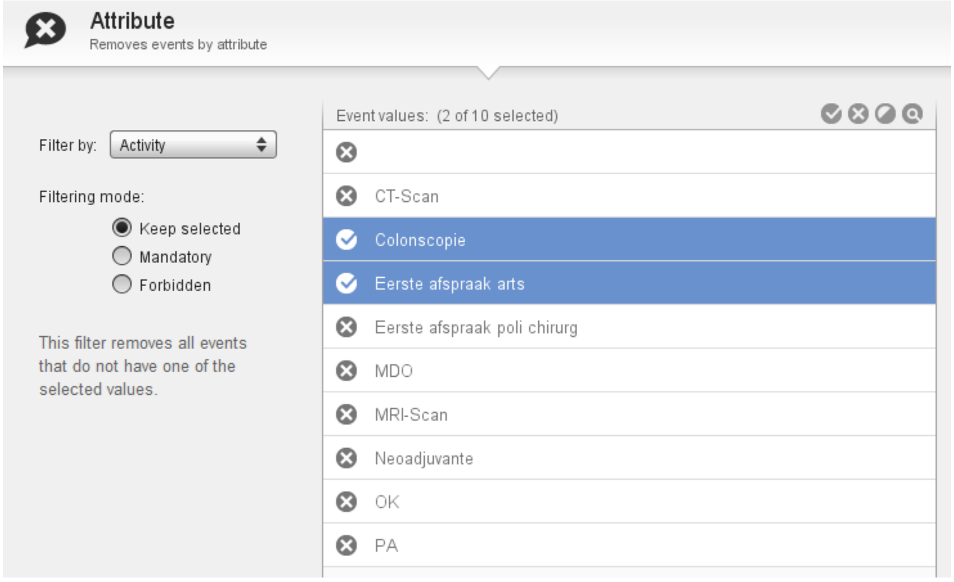
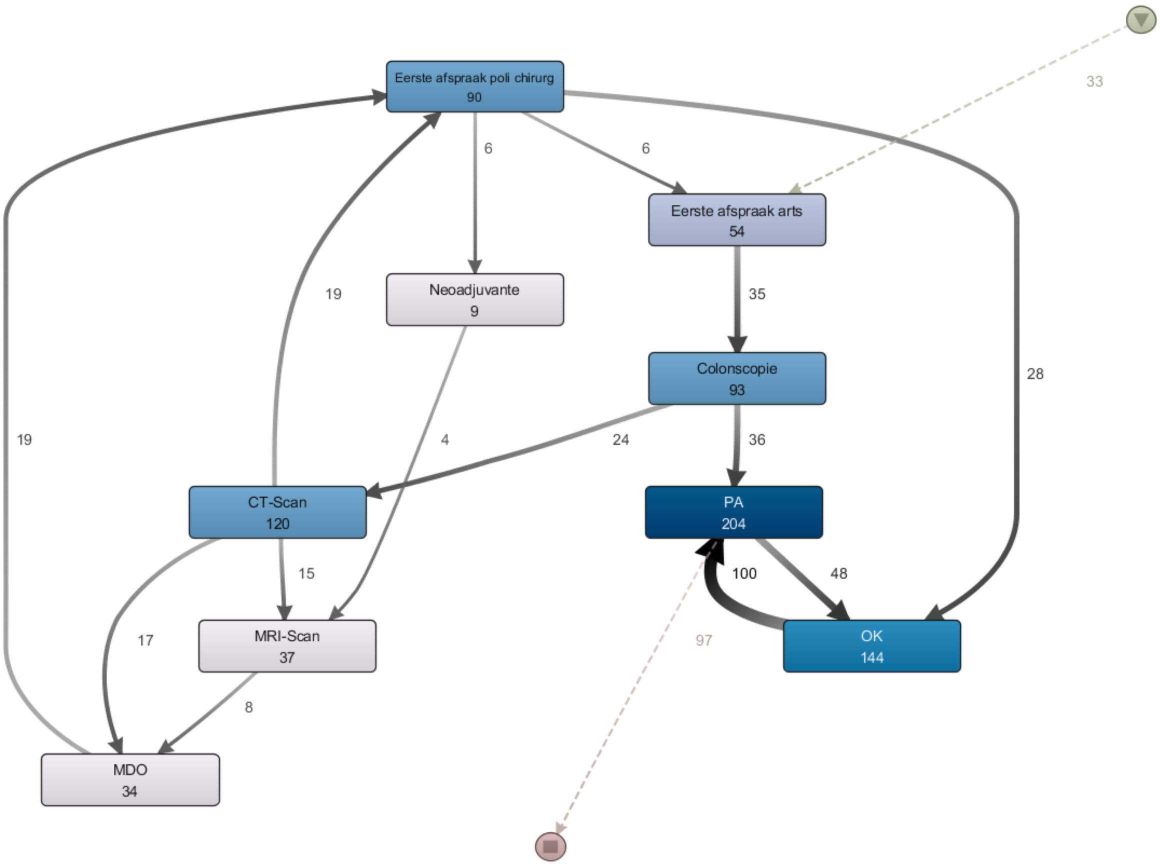


Figure 39: An attribute filter in DISCO to create a process model



4. Perform Analyses

4.1 Process model analyses

The analyses of the datasets were made with the focus on the two main questions. The manager of the department helped with creating the analyses by explaining the execution of (certain) activities. The following paragraph will describe the results of both analyzed questions.

Creating an overview of the number of patients who were treated in the year 2014:

As mentioned before, this dataset is a combination of three different datasets. Each dataset contains several treated patients from the year 2014. The three data sets are used in the care cycle that patients pass through, whereby several patients occur in multiple datasets. To find out what the unique number of patients in the three datasets is together, a function was used in DISCO which made the unique number of patients visible. This has shown that 2761 patients were treated for colorectal cancer in 2014. Also, in the EHR-system, PA, and DSCA dataset there are respectively: 510, 2552, and 280 unique patients.

Furthermore, an overlap of unique patients was seen with the following amounts, namely: the EHR-system and PA have an overlap of 390 patients, the EHR- system and DSCA have an overlap of 223 patients, and PA and DSCA have an overlap of 357 patients.

Creating overviews of turnaround times of the most followed paths:

Selecting the appropriate process steps was adjusted to the relevant activities per turnaround time.

The first turnaround time we looked at was from first doctor visit to colonoscopy.

The created model showed that a total of 138 patients went through this process. However, 128 patients first underwent a doctor visit and 10 patients first underwent a colonoscopy. Furthermore, 132 patients had a colonoscopy after their first doctor visit.

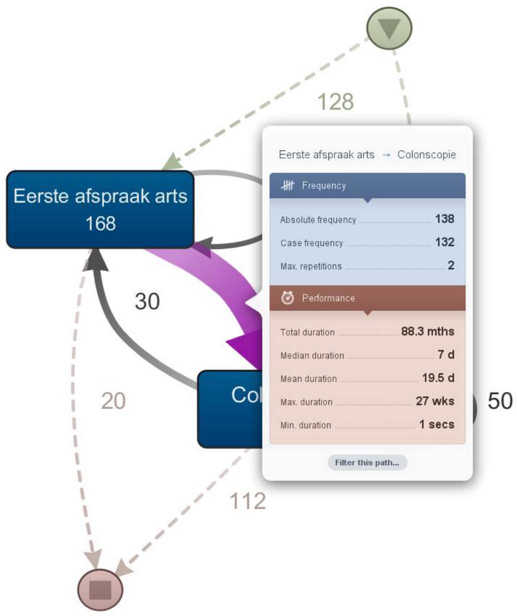


Figure 40: Number and turnaround times of patients between First Doctor Visit and Colonoscopy

In total this process lasts approximately 20 days per patient. Which indicates that it takes a patient about 20 days to have a colonoscopy from their first doctor visit.

The next analyzed process was from Colonoscopy to Pathological Examination.

In this process a total of 137 patients were involved, where 131 patients first had a colonoscopy and 6 patients a pathological examination after their first doctor visit. Although some patients first got a pathological examination, according to the data, afterwards also had a colonoscopy.

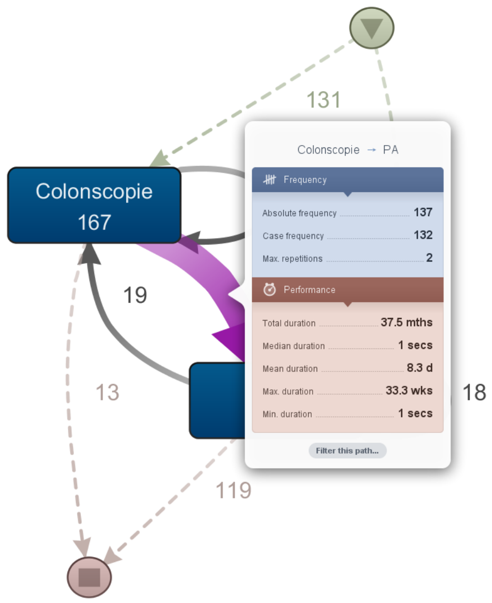


Figure 41: Number and turnaround times of patients between Colonoscopy and Pathological Examination

In total this process lasts approximately 8 days per patient. Which indicates that it takes a patient about 8 days to have a pathological examination after a colonoscopy.

The next analyzed process was from Colonoscopy to a CT-Scan.

In total 244 underwent this process, of which 220 patients had a colonoscopy and afterwards a CT-Scan and 24 patients first had a CT-Scan and then a colonoscopy. From a CT-Scan 50 patients went back to the colonoscopy, which could indicate that patients from colonoscopy to CT-Scan went back to colonoscopy or patients from CT-Scan went back to colonoscopy to CT-Scan and again back to colonoscopy. As can be seen, it is a very devious execution of this process.

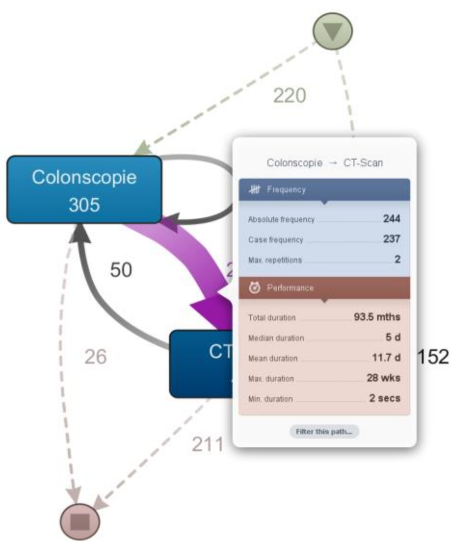


Figure 42: Number and turnaround times of patients between Colonoscopy and CT-Scan

Furthermore, it takes around the 12 days for a patient to get through this process.

The next analyzed process was from Colonoscopy to a MRI-Scan.

In this process 118 patients were treated, where 110 patients went straight to colonoscopy and 8 patients first went to MRI-Scan. Also, can be seen that only 111 patients from the 118 are ending this process. This could indicate that the remaining patients of 7 ended in a loop in this process or that this process has not been ended in the year 2014.

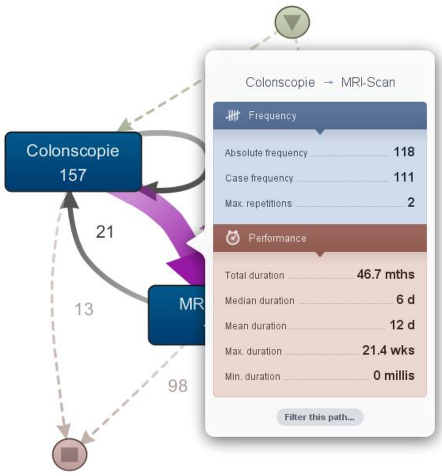


Figure 43: Number and turnaround times of patients between Colonoscopy and MRI-Scan

In this process a patient took approximately 12 days to go through it.

The next process what was analyzed was Colonoscopy to a Multidisciplinary Consultation.

All the patients, namely 142, in this process directly went to colonoscopy and from there to the multidisciplinary consultation. In the activity colonoscopy 10 patients occurred to be in a loop, which means that those patients had a second colonoscopy (according to the data).

Figure 44: Number and turnaround times of patients between Colonoscopy and Multidisciplinary Consultation

Also, around the 11 days was needed to treat a patient in this process.

The next process which was analyzed was Multidisciplinary Consultation to a First Surgeon Visit.

In the analyzed of this process it became visible that this was a straight forward process. A total of 145 patients started this process and the same amount went through the end.

Figure 45: Number and turnaround times of patients between Multidisciplinary Consultation and First Surgeon Visit

For one patient, it lasted 10 days to reach the end of a first surgeon visit-activity.

The next and last analyzed process was Multidisciplinary Consultation to a Surgery.

In the last analyzed process, a total of 163 patients were involved. From these numbers, 2 patients first went directly to Surgery and then to the multidisciplinary consultation. This flow is not possible because in a multidisciplinary consultation it is discussed what treatment is suitable for a patient and its condition.

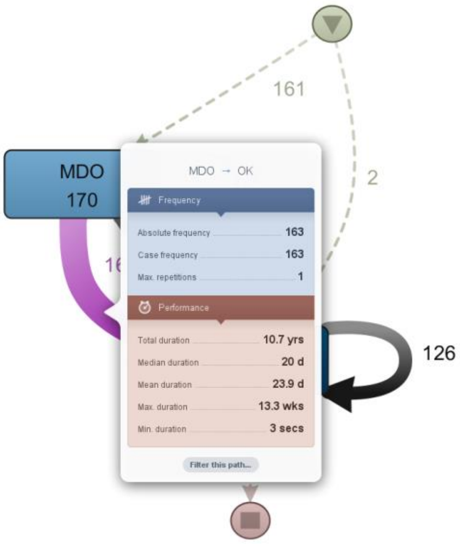


Figure 46: Number and turnaround times of patients between Multidisciplinary Consultation and Surgery

One patient needed approximately 24 days to go through this process.

5. Discussing and Validating

5.1 Report the analyses

The created process models were discussed using a power point presentation. In this report, there are several snapshots used. We discussed the information obtained from Process Mining with content experts. In this manner, the information from the analyses could be interpreted in a proper form.

With this research, we could indicate that a view or model of processes can be created in a shorter time, which in time will save money. In a few days, process models can be modelled with large amounts of data. Due to this, frequent measurements of throughput times can be created. Also, necessary adjustments can be made in a timely manner.

This all is possible with the help of filters in a Process Mining application. These filters helped, in this research, to answer specific questions about the Colorectal processes.

5.2 Transferring the results

The results of the throughput times of the clinical pathway were presented to the internist oncologist, surgeon, stomach intestine liver specialist, and business analyst. Then, the results were communicated to the doctors in a multidisciplinary meeting. The doctors reacted critical, since the results and analyses possibilities should be available within the current EHR-system.

The eight of the nine measure points could be measured with more data in a shorter period. Normally it took the hospital months to create throughput times, which now can be reduced to a days' work with a larger sample, since all the data of the patients are used for the analyses.

7 VALIDATION

This chapter will elaborate on the usability of the Care Mining Method. The usability is tested by the modeler and a medical expert. In the first paragraph the validation is done by the modeler of this research and in the second paragraph the validation is done by a medical expert of the hospital were the cases were conducted.

7.1 VALIDATION BY THE MODELER

The modeler of the Care Mining method validated the method by evaluating each created step. After the method was applied on both cases the modeler determined per step to which extend the degree of implementation of the method was. The validation has been elaborated in Table 4.

To analyze the results of both cases the Process Mining application DISCO was used. DISCO was chosen to be used, because it is quick and easy to install, and quick and easy to learn and to apply with large and complex datasets.

The modeler had not worked with DISCO before and had to learn how to work with it. This was done by reading about the functions of the application and by applying random test cases and investigate what the results are and how to interpret these. DISCO was installed in one day and it took the modeler one week to be able to understand and apply the application.

In general, the modeler has six years of experience with modeling processes. In these six years, the modeler has also worked with some tools, to create a process model (i.e. Microsoft Visio).

Care Mining Method	Degree of applicability	Argumentation
1. Select a medical condition		
1.1 Gain knowledge about the medical condition	Easy	This step could be applied quickly, because a lot of information could be found on websites and during interviews.
1.2 Define the scope of the research	Difficult	Defining the scope of a research was not always easy to define. Hospitals can have different starting activities, depending on the different departments and different treatments of patients. In some cases, several meetings were needed to be able to define a suitable starting-and ending activity. To define the scope, hospitals had to have clear question about what needed to be researched. According to the questions, a scope could be defined faster. In some cases, hospitals do not have it clear about what needs to be researched. Most of the times hospitals need help in created research questions.
2. Define Care Delivery Value Chain (CDVC)		
2.1 Gain information about care cycle activities	Simple	This step was easy to define. During an interview with medical specialists, information about activities in care cycles can be simply gained.

2.2 Define performed activities	easy	Defining performed activities is easy to describe. This step can be performed during an interview with medical specialists. In these interviews the operation and application of these activities can be asked for.
3. Develop process-maps		
3.1 Identify and define dataset	Normal	This step was not very easy to apply. The reason was that in most cases it was clear about what information was needed, but that it took time for the medical specialists to explain and reach the business analysts to discuss from where data exports was needed.
3.2 Obtain patient-related data	Simple	Obtaining patient- related data was easy to apply, because in this step the researchers got the datasets from the business analysts.
3.3 Log preparation	Difficult	Applying this step took some time. Most datasets had to be transformed into an extension what could be read by a Process Mining tool. This process could take some hours.
3.4 Log inspection	Difficult	The log inspection- step took some time in performance. To perform a log inspection, which is according to the research questions, interviews with business analysts were needed. They could explain and describe the data in the different columns of the datasets. In most cases the researchers received more data then was necessary. Therefore, in all the cases some columns had to be deleted from the datasets.
3.5 Create process model	Normal	This step was easy to be performed. In this step the dataset is ready to be imported into a Process Mining tool. After the dataset was imported, with the help of the functions of the tool, process models could be created in minutes.
4. Perform Analyses		
4.1 Process model analyses	Normal	Creating analyses of created process models was quite easy to do. This means that the researcher was able to describe the created process models. The difficult part of this step was determining what the reasons of some different deviant patterns in the models are. In some cases, the medical specialists or business analysts were not able to explain some activity flows.
5. Discussion and Validation		
5.1 Report the analyses	Simple	In all the performed cases, the results were reported in a PowerPoint presentation. The researchers chose this type to report the results, because in this case the models could be explained clearer.
5.2 Transferring the results	Simple	The created models presented in a power point presentation made it more convenient to transfer and discuss the results with medical specialists.

Table 4: Validation of the Care Mining Method

We could analyze that two steps were *easy*, four were *simple*, two were *normal*, and three were *difficult* to be performed in practice. The difficult steps had as reason that the combination of the knowledge of the modeler and the knowledge of the involved medical specialist lacked to give a quick response. This resulted in more meetings or discussions to gain data or information or to be able to analyze results.

7.2 VALIDATION BY AN EXPERT

To be able to determine whether the results of the cases and the performance of the method were as expected, an interview was held with a medical specialist. The validation of the method and results was done using four open questions, which will be demonstrated in the next paragraphs.

This validation interview was held with a physician from the same hospital where aforementioned cases were performed in, namely Amhos. The following will elaborate on and describe the answers of the physician.

1. Are the results as expected?

The results were as shocking as surprising. In the Colorectal case, for example, the created process models made it visible that the activity-flows were different in practice than in theory. The physician mentioned that it is helpful to have depictions of the processes, which will lead to more awareness and insights into the current processes.

2. Is the method interesting?

The physician indicated that this method is very interesting to apply in the healthcare sector. With this method hospitals are able to manage their processes faster and easier than their current approach. Currently, process models are created manually and are time consuming. Therefore, the process models are not created for every process, but only for the once where the physicians think that certain improvements can be made. For example, normally the medical specialists need to wait months for turnaround times of a patient, which now can be generated in a few weeks. This was a major asset of this method.

3. Are there more questions about the method?

There were a few more questions the physician had about the usefulness of the method. The physician praised the way turnaround times are reported in this way. A question about this was if it is possible to answer ad hoc questions about certain turnaround times. We answered that this is possible with a Process Mining tool. With such a tool, any question could be answered with the right transformed and imported dataset in it.

4. Is the method applicable in the organization?

This question was a very important one for the physician, because of cost issues. The physician mentioned that financially it would cost the hospital more if the method is applied by an external party, which was the case now. In a preferable case, the hospital has its own personnel who can apply and interpret this method.

To achieve this goal, a small trainings session could be organized for the medical personnel. In this session, the personnel can learn about the explanations of the method steps and the usage of a Process Mining tool.

8 DISCUSSION

This chapter will elaborate on the discussion of this research. The discussion is divided into main findings, challenges and limitations and future work.

8.1 MAIN FINDINGS

In this thesis four sub-questions were addressed that we wanted to answer. In this paragraph the main findings of this research will be described and each sub- and research question will be answered and.

8.1.1 MAIN FINDINGS

1. What went wrong?

During the performance of this investigation there were a few elements that did not progressed as planned. In the next paragraph these elements are discussed.

Creating the Care Mining Method

A theoretical study was performed to create the Care Mining Method, to combine two existing methods into one. After the first version of the method was created, it was discussed with two experts, namely: a medical informatician and a process miner. From this discussion, we concluded that the method was not efficient enough. Because, the method consisted of too many steps and eventually it would cost the hospital more time and money to perform all the steps. Finally, to create a method which could be used in a real case took significant more time than was stated in the predefined planning.

Obtaining cases

To test the performance of the Care Mining Method, hospital cases were required. Meetings with several hospitals were held to talk about this research and to determine whether and how the method could be performed and be of benefit of the concerned hospital. Unfortunately, due to a lack in awareness of the importance of analyzing care processes, research time restrictions, and or budget limitations more cases could not be acquired. Finally, two cases could be obtained for this research, instead of a minimum of three.

Determining a research scope

From both cases, we could find that when we started to perform the method, defining a scope was quite difficult. The problem was that hospital personnel could acknowledge that somethings are not going the way it supposed to be, but lack in identifying or pointing out the bottlenecks. Thus, to determine in what area the focus of the research should be, we had to schedule more interviews and meetings than was planned.

Also, more meetings were necessary to validate the temporary findings of the research and to be able to create proper interpretations. All these interviews and meetings lead to a longer investigation per case, regarding the time.

Validating the Care Mining Method

After each case was performed meetings were held to discuss the results of the concerned investigation with the involved people. Also, the method needed to be validated generally and therefore we had to interview involved people of one of the cases from the hospital. Scheduling an interview appeared to be difficult, because a lot of time had passed since the completion of the cases and the time of planning a final interview. The hospital assumed that everything had already been completed and did not count on another appointment. Therefore, it took longer before an appointment could be scheduled.

2. How can we improve the future?

To learn from all the above-mentioned elements of the research that did not went the way which was planned in advance, we will discuss how those elements can be performed in the future.

Creating the Care Mining Method

Creating the Care Mining Method needed a significant amount of time than was planned. For in the future, experts should be involved earlier in the processes of creating a method. Thus, certain improvement points will be discovered, discussed, and adjusted much faster.

Obtaining cases

Obtaining a case was difficult because we were too awaiting and hopeful that a hospital or department of a hospital would like to cooperate right away. We had to approach other hospitals during discussions with certain hospitals. For the next time, we need to approach several hospitals at once. In this way one avoids losing too much time in waiting for response and the need for re-approaching other hospitals.

Hence, it is effective to make contacts at the same time, because time is needed to build trust and to be able cooperate in an investigation.

Determining a research scope

It must be considered that determining a scope requires a certain amount of time. This time is needed for the researchers to create appropriate interpretations of the data and to be able to communicate the analyzed data with the hospital. So, when a planning is being made more time should be scheduled for determining the scope.

What is important is that it should be communicated with the hospital what the total amount of research time will be. In this way, it is more clear for the hospital to determine whether they have the space for such a research. This part refers to the first point **Creating the Care Mining Method**.

Validating the Care Mining Method

Validating the Care Mining Method required more time than was planned. What can be learned from this point, is that the researchers should plan in advance that the method needs to be validated. When creating a planning, this validation part should be scheduled beforehand with the concerned hospital.

3. Is this research valid?

To check whether this research is valid, the validity criteria of Yin (2009) were performed as described in section **2.2. Validity** of this research. The construct validity describes the identification of the right operational measures (Yin, 2009) and was performed with two experts, the external party Zuiver ICT and the concerned hospital. With the two experts, the method and its performance was validated. Validating the method was executed with meetings. Internal validity describes to identify a causal relationship (Yin, 2009) was done with Zuiver ICT and with the concerned hospital. With Zuiver ICT the obtained cases were validated on its usefulness for the method and with the hospital meetings were held with involved people of the cases. The pre-arranged expectations of the hospital served as a measure. From the cases, we had received a process flow of all processes, which served as a benchmark. Finally, empirical reliability describes that same results should be the outcome when performing the same research. This part could not be performed due to a lack of time for this research.

4. Is this research useful?

This research is useful, because a method is proposed about how an overview can be created in an efficient and effective way, concerning the performance time of care activities. There are existing researches which proposes a method to identify and analyze hospital processes (Witowski, Higgins, Warner, Sherman, & Kaplan, 2013; Kaplan, 2014), but they do not combine it with a Process Mining tool. But, there is one research of Mans, Aalst, & Vanwersch (2015) which introduces Process Mining in the Health Care sector. That research was performed in other hospitals than is described in this thesis. Also, this thesis can be distinguished from the research from Mans, Aalst, & Vanwersch (2015) by the Care Mining Method. The Care Mining Method describes the whole process of defining data till analyzing the created process-models. Mans, Aalst, & Vanwersch (2015) propose a, so called, Healthcare Reference Model. This model describes how hospital event data can be easily located and extracted.

5. What can the researchers learn from the research methods?

The researchers from this thesis can learn that it is possible to combine the theories of Value-Based Health Care and Process Mining. Despite the extra time which was needed to perform certain steps in the combined method, Care Mining Method, the actual extraction of data and the creation of a process-model made it an effective and efficient method.

8.1.2 SUB-QUESTIONS AND RESEARCH-QUESTION

SQ1: *"What is Value-Based Health Care and how are costs related to it?"*

Value-Based Health Care is a method that aims at maximizing patient care and reducing healthcare costs. To be able to maximize patient care or reduce healthcare costs, analyses about the performed processes must be executed. These analyses are made on performed health care activities and is supported by patient and physician-relevant indicators (such as survival, complications and quality of life) and process indicators and costs. These activities together form processes that are present in a care cycle. The analyses are made to research to what extent predefined objectives are achieved. After the analyses are made, medical specialist (e.g. oncologist) together with business analysts can create improvements based on the results of the analyses.

The costs or performance time of each activity in a care cycle identifies the treatment period of one patient and helps in creating process models. Mapping the costs is done with a Process Mining tool.

SQ2: *"What is known about Process Mining in Health Care?"*

In theoretical findings, we found that there is little known about Process Mining in Health Care. The main reason is that the Health Care sector needs to build more awareness about the importance and benefits of analyzing their processes. When the awareness and knowledge of medical specialists is increased, more studies about process analyses will be allowed to be performed. Also, time and finances are required to perform Process Mining studies. The studies with Process Mining that were performed in Health Care were executed by technical skilled researchers outside the hospital, rather than personnel from the hospital itself.

SQ3: *"How can Process Mining support VBHC for identifying and improving costs?"*
Process Mining in the Health Care sector helps in identifying and analyzing health care processes by using the Process Mining Process, including six steps, and the Time-Driven Activity-Based method, including seven steps. The Process Diagnosis Method is used for the support of Process Mining, to create process models. The Time-Drive-Activity-Based Method is used to support VBHC for identifying costs, to form a base (i.e. determining a medical condition) and structure (i.e. determining the research scope) for the analyses. A combination of these two methods resulted in a new method, which lead to the creation of analyses of hospital processes. Thus, providing a solid basis to manage and improve processes within hospitals.

SQ4: *"How can the support of Process Mining within VBHC be put in practice?"*
The aforementioned built method is created to be able to put both methods of Process Mining and VBHC together in practice. The method has been put in practice and tested in two different hospital cases in The Netherlands. To use the method interviews with hospital personnel is required to discuss what needs to be researched and to be able to obtain and build knowledge about the research area. Hospital personnel should be people who work with patients in practice (i.e. medical specialists), to provide practical information about care processes, and people who work with the processes (i.e. business analysts), to provide operational information about the care processes. The latter also helps in providing the datasets with executed patient data. After data is acquired and process models are created and analyzed, the results must be discussed with the same group of people who are involved in the investigation. Furthermore, patient-related data from the hospital's EHR information system is required, to create process models that represent the reality.

RQ: *"How can Process Mining strengthen Value-Based Health Care?"*
It has been shown that the method of Process Mining can be well integrated with the method of VBHC. The results of the performed cases showed that by using Process Mining with VBHC the care processes could be modeled in a couple of hours in one day, instead of in weeks. This is an important advantage of this method, referring to the limited budgets and rising costs in the Health Care sector. Furthermore, Process Mining techniques helps in creating a clear overview of care processes, which in the end supports in increasing value in care. The latter will be accomplished when the process overviews are mapped with predefined objectives and process maps.

Concluding, Process Mining strengthens VBHC due to the following, namely: 1) decreasing the performance time of analyzing care processes, 2) more visual and distinct process models can be created, and 3) the process models can now be maintained more accurately and effectively.

8.2 CHALLENGES AND LIMITATIONS

The described method in this research was not performed by a specialist from a hospital and could be done in a future research. This is a limitation for this research, because in practice the hospital personnel itself should learn how to integrate and use the Care Mining method. It is possible for a hospital to hire technical skilled people to execute the Care Mining method, but then the financial aspects can be jeopardized (due to limited budgets).

Process Mining requires technical knowledge to be performed in practice (e.g. data transformation need to be done to analyze the dataset from the information systems). For medical personnel to use the Care Mining method it is required that they take trainings. The trainings should be about the Process Mining part of the Care Mining method and not Process Mining in general, which should shorten the training days and should result in a more comprehensible training. How much time is required for a training should be discussed with the hospital. These trainings can be seen as a challenge and limitation. It is a challenge, because medical personnel will be trained into a technical field where they initially were not educated in. The limitation of the trainings is in the fact that time is needed in order to give the trainings to medical personnel, which most of the time is given to treat patients.

A challenge to use the method is that time, and linked to that finances, are necessary. Time is required for the interviews and discussion, and for the trainings for personnel. Combined with the time additional finances are needed to train personnel and perform the Care Mining method. Advised is that, a research team is formed to perform the research, which can benefit the structure and results of the research.

Also, to use the techniques of Process Mining a tool is required. There are tools which can be used for free, when working with a maximum of data records. However, in the case of hospitals there are too many data records that a tool must be purchased for usage.

Another limitation is that, exact times of all the activities can sometimes not be notated. This is because in some cases a doctor does not always record data directly in the system, as this may be a few days or a week later than the actual one. Therefore, field research is required to investigate the actual execution and record of care processes in an information system.

Finally, to the best of our knowledge this is the first method that focuses on the integration of VBHC and Process Mining in the Health Care sector. The combination of VBHC aims to analyze processes in a step-by-step manner and with the involvement of hospital personnel.

8.3 FURTHER WORK

During and after performing this research it became clear that some other research topics were triggered. These topics can be investigated further in a future research for more amplification.

Due to time issues, this research could not be implemented in a hospital. Further research in practice is needed to investigate the applicability of the method when performed by hospital personnel itself and the time that is required to train personnel.

For the sake of the simplicity and use of a Process Mining tool for the hospital, DISCO was chosen. This tool was not only simple to understand, but also low in cost in terms of purchase. In the future, another Process Mining tool could be used to investigate and compare the results of medical processes.

This research was focused on processes, which involved patients and their throughput times. These processes can be expanded by analyzing the working days and the resources of medical specialists. This leads to another expansion of this research, which involves adding financial costs to the performed activities in a care cycle to determine the financial costs of throughput times. In other words, a third expansion concerning financial costs, is calculating the costs of a whole care cycle with all the used facilities.

Also, the maintenance of the created process models could be investigated further. This research highlights the preservation of implemented improvements and should prevent the occurrence of certain bottlenecks. Maintaining process models can be executed by business analysts and requires a Process Mining tool with a display, preferable a dashboard, from where the performed care processes could be managed.

9 SUMMARY

This thesis was focused on reducing costs by inspecting the flow and time of care processes. Due to budget limitations and increasing costs in hospitals, further insights and research was necessary to identify how and where costs could be reduced.

A number of studies have been performed to create transparency in care processes. These studies were mostly directed towards making the finances of hospitals transparent and researching quality management. The quality of hospital services is dependent on suitable and efficient performances of care processes. These processes are several activities together to diagnose, treat, and prevent diseases for the improvement of a patient's health. The execution of these processes is done by different resources (e.g. physicians, technical specialists, nurses, etc.).

Analyzing and improving health care processes is a challenging task, due to its dynamic, complex, ad-hoc, and multidisciplinary performances. Other studies have shown that it is not easy to improve the processes and a number of challenges could occur. However, improving these processes could result in a high impact on the quality of life of patients. Therefore, a method was created in this thesis to facilitate the analyses.

The purpose of this thesis was to analyze whether Process Mining as a method could be combined with a method of VBHC. The reason to research the integration of these two methods was to investigate if the created method could influence the effectivity and efficiency of the analyses of care processes. To be able to create a method, a literature study and case studies were performed. The literature study on both methods formed a base to create a new method.

In the literature study, we found that limited studies investigate the way in which healthcare value can be highlighted and increased with the help of Process Mining. The studies performed in The Netherlands used Process Mining as a tool to analyze hospital data. No steps or phases were presented to indicate the acquisition of patient-related data and the communication with hospital personnel. Furthermore, a method called Time-Driven Activity-Based Costing was found in literature, which explains a step-by-step approach to identify, obtain, and analyze care processes, with a focus on increasing value in health care. This method was applied in case studies in USA and proved to be an effective method in the health care sector. Therefore, this method was chosen for this thesis to be applied in Dutch cases. For the sake of efficiency of the performance time of the TD-ABC method, a Process Mining Process method was used. The combination of TD-ABC and the Process Mining Process method resulted in a new method called Care Mining Method.

The Care Mining method was put in practice in two hospital cases. More conversations and meetings with other cases were held, but due to a lack in awareness of the importance of analyzing care processes, research time restrictions, and or budget limitations more cases could not be acquired. But this can be performed in a further research.

Furthermore, after obtaining patient-related data the process models were created in a few hours in accordance to the prepared questions. Particularly this part of the method was a major improvement, because with the TD-ABC method the process models had the be created manually. Finally, after analyzing the process models we could find that care processes are performed different than was predicted or expected. With these models the managers and or business analysts could set up another project to discuss if and how certain care processes could be improved or changed.

10 REFERENCES

Aalst, W. van der. (2011). *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Springer Publishing Company. http://doi.org/10.1007/978-3-642-19345-3

Aalst, W. van der, Adriansyah, A., Medeiros, A. K. A. de, Arcieri, F., Baier, T., Blickle, T., … Wynn, M. (2012). Process Mining Manifesto, 169–194.

Agrawal, R., Gunopulos, D., & Leymann, F. (1998). Mining Process Models from Workow Logs. *6th International Conference on Extending Database Technology*, 467–483. http://doi.org/10.1007/BFb0101003

Bhattacherjee, A. (2012). *SOCIAL SCIENCE RESEARCH: Principles, Methods, and Practices* (2nd ed.). Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.

Bloomberg. (2013). Biggest Rise in Health-Care Cost Relative to Income: Countries. Retrieved from http://www.bloomberg.com/visual-data/best-and-worst/biggest-rise-in-health-care-cost-relative-to-income-countries

Boer & Croon. (2008). *Naar waardecreatie in de zorg - Een vertaling van "Redefining Health Care" naar de Nederlandse situatie*.

Bozkaya, M., Gabriels, J., & Werf, J. M. Van Der. (2009). Process Diagnostics: A Method Based on Process Mining. *International Conference on Information, Process, and Knowledge Management*, (1), 22–27. http://doi.org/10.1109/eKNOW.2009.29

Caron, F., Vanthienen, J., Vanhaecht, K., Van Limbergen, E., Deweerdt, J., & Baesens, B. (2014). A process mining-based investigation of adverse events in care processes. *Health Information Management Journal*, *43*(1), 16–25. http://doi.org/10.12826/18333575.2013.0013.Caron

Centraal Bureau voor de Statistiek. (2014). Uitgaven als percentage van het bbp. Retrieved from http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=81628NED&D1=93&D2=a&HDR=G1&STB=T&VW=T

Charles, D., Furukawa, M., & Hufstader, M. (2012). Electronic health record systems and intent to attest to meaningful use among non-federal acute care hospitals in the United States: 2008-2011. *ONC Data Brief*, (1), 1–7.

Chen, J., Ou, L., & Hollis, S. J. (2013). A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Serv Res*, *13*(1), 211. http://doi.org/10.1186/1472-6963-13-211

Cook, J. E., & Wolf, A. L. (1998). Discovering models of software processes from event-based data. *ACM Transactions on Software Engineering and Methodology*, *7*(3), 215–249. http://doi.org/10.1145/287000.287001

Cooper, R., & Kaplan, R. S. (1999). *The design of cost management systems : text and cases* (2nd ed.). Upper Saddle River [etc.] : Prentice Hall.

Corbey, M. (2008). Time-Driven Activity-Based Costing, (November), 477–484.

DICA. (2016). Dutch Surgical Colorectal Audit (DSCA). Retrieved from https://www.dica.nl/dsca/home

Dol, C., Raap, M., & Lageman, P. (2014). De meerwaarde van Value Based Health Care, p. 3/24.

Gartner. (2016). Automated Business Process Discovery (ABPD). Retrieved from http://www.gartner.com/it-glossary/automated-business-process-discovery-abpd

Ginter, P. M., Duncan, W. J., & Swayne, L. E. (2013). *Strategic Management of Healthcare Organizations* (7th ed.). United Kinkdom: John Wiley & Sons.

Goedee, A., Suttorp, O., & Lieshout, H. van. (2014). *Met aandacht voor zorg Jaarverslag 2014*. Retrieved from https://www.amphia.nl/~/media/files/amphia_jaarverslag2014.ashx?la=nl-nl

Hagen, M. van. (2015). Enabling Value Based Health Care. Retrieved from http://www.slideshare.net/matthijsvanhagen/4-walsteijnedifecs-enabling-value-based-healthcare-2015-04-09-v3

Hennrikus, W. P., Waters, P. M., Bae, D. S., Virk, S. S., & Shah, A. S. (2012). Inside the value revolution at Children's Hospital Boston: time- driven activity-based costing in orthopedic surgery. *The Harvard Orthopaedic Journal*, *14*(December).

Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science Research in Information Systems. *MIS Quarterly*, *28*(1), 75–105.

Hilsenrath, P., Eakin, C., & Fischer, K. (2015). Price-Transparency and Cost Accounting: Challenges for Health Care Organizations in the Consumer-Driven Era. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 1–5.

Hinkin, T. . R., Tracey, J. . B., & Enz, C. A. (1997). Scale construction: Developing reliable and valid measurement instruments. *Journal of Hospitality and Tourism Research*, *21*(1), 100–120. http://doi.org/10.1177/109634809702100108

James, B. C. (1989). *QUALITY MANAGEMENT FOR HEALTH CARE DELIVERY* (Catalog No). The Hospital Research and Educational Trust.

Jeffries, R., & Desurviret, H. (1992). Usability Testing vs . Heuristic Evaluation : Was there a contest ?, (October), 39–41.

Kaplan, R. S. Introduction to Time-Driven Activity- Based Costing in Health Care (2013).

Kaplan, R. S., & Anderson, S. R. (2003). Time-Driven Activity-Based Costing. *Harvard Business Review*, *82*(November), 131–138. http://doi.org/10.2139/ssrn.485443

Kaplan, R. S., & Anderson, S. R. (2004). Time-Driven Activity- Based Costing. *Harvard Business Review*, *82*(11), 131–138.

Kaplan, R. S., & Anderson, S. R. (2007). The innovation of time-driven activity-based costing. *Journal of Cost Management*, *21*(2), 5–15.

Kaplan, R. S., & Porter, M. E. (2011). How to Solve The Cost Crisis In Health Care, (September).

Kaplan, R. S., Witkowski, M., Abbott, M., Guzman, A. B., Higgans, L. D., Meara, J. G., … Feeley, T. W. (2014). Using Time-Driven Activity- Based Costing to Identify Value Improvement Opportunities in Healthcare. *Journal of Healthcare Management*, *59*(6), 399–413.

Kirchmer, M., Laengle, S., & Masias, V. (2013). Transparancy-Driven Business Process Management in Healthcare Settings. *IEEE Technology and Society Magazine*, *32*(4), 14–16.

Koomans, M. (2014). Does Value Based Healthcare make care cheaper and better? Retrieved from http://www.skipr.nl/blogs/id1803-maakt-value-based-healthcare-zorg-goedkoper-en-beter.html

KPMG. (2013). *Measuring the Value of Healthcare Delivery: Cutting through complexity*.

Kuenen, J. W., Mohr, R., Larsson, S., & Leeuwen, W. van. (2011). *Zorg voor waarde*.

Kunst, P. E. J., Lemmink, J. G. A. M., & Prins, R. (1996). Interrelationships Between Key Parameters of Service Quality. A Study to Explore Best Practices in the Service Sector in EU Member States, Report for the European Commission.

Kunst, P., & Lemmink, J. (2000). Quality management and business performance in hospitals: A search for success parameters. *Total Quality Management*, *11*(8), 1123–1133. http://doi.org/10.1080/095441200440377

Lappi, M. D., Flath-Sporn, S. J., Heald, R., Kim, D. C., & Meara, J. G. (2015). Increasing Value in Plagiocephaly Care: A Time-Driven Activity-Based Costing Pilot Study. *Annals of Plastic Surgery*, *74*, 672–676.

Lenz, R., & Reichert, M. (2007). IT support for healthcare processes - premises, challenges, perspectives. *Data and Knowledge Engineering*, *61*(1), 39–58. http://doi.org/10.1016/j.datak.2006.04.007

Mans, R. S., Aalst, W. M. P. Van Der, & Vanwersch, R. J. B. (2015). *Process Mining in Healthcare Evaluating and Exploiting Operational Healthcare Processes*. Springer International Publishing AG Switzerland.

Mans, R. S., Schonenberg, M. H., Song, M., & van der Aalst, W. M. P. (2008). Application of Process Mining in Healthcare – A Case Study in a Dutch Hospital. *Biomedical Engineering Systems and Technologies - Communications in Computer and Information Science*, *25*, 425–438. http://doi.org/10.1007/978-3-540-92219-3_32

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Elsevier*, 251–266.

Marianne Heins, Spronk, I., Jong, J. de, Ho, V., Brink, M., & Korevaar, J. (2015). *Variatie tussen ziekenhuizen in de behandeling van vijf soorten kanker. Een verkennend onderzoek naar aanknopingspunten voor verbetering van zorg*. Retrieved from https://www.nivel.nl/sites/default/files/bestanden/behandeling-kanker-ziekenhuizen-mammacarcinoom.pdf

Mcclellan, M., & Mcclellan, M. (2011). Reforming Payments to Healthcare Providers : The Key to Slowing Healthcare Cost Growth While Improving Quality? Reforming Providers : Cost Growth While Improving Quality? *The Journal of Economic Perspectives*, *25*(2), 69–92.

McLaughlin, N., Setlur, N. P., Kaplan, A. L., & Mahajan, A. (2014). Time-driven activity-based costing: a driver for provider engagement in costing activities and redesign initiatives. *Neurosurg Focus*, *37*(November), 1–9.

Medeiros, A. de, Karla, A., van der Aalst, W., & Pedrinaci, C. (2008). Semantic process mining tools: Core building blocks. *Proceedings of the 16th European Conference on Information Systems (ECIS '08)*, 1953–1964. http://doi.org/http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.141

Ministerie van Volksgezondheid, W. en S. (2016). Jaarverslagen Zorg. Retrieved from https://www.desan.nl/net/DoSearch/Search.aspx

Monroy, C. R., Nasiri, A., & Peláez, M. Á. (2014). Annals of Industrial Engineering 2012. *Springer*, 11–18. http://doi.org/10.1007/978-1-4471-5349-8

Mueller-Wickop, N., & Schultz, M. (2013). ERP event log preprocessing: Timestamps vs. accounting logic, *7939 LNCS*, 105–119. http://doi.org/10.1007/978-3-642-38827-9_8

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *Jama*, *309*(13), 1351–1352. http://doi.org/10.1001/jama.2013.393

Mutze, B., & Ierland, D. van. (2007). Time Driven ABC: nieuw in Nederland?, 42–45.

Petrova, M., Dale, J., & Fulford, B. (2006). Values-based practice in primary care: Easing the tensions between individual values, ethical principles and best evidence. *British Journal of General Practice*, *56*(530), 703–709. http://doi.org/333/7560/173 [pii]\r10.1136/bmj.333.7560.173 [doi]

Porter, M. E. (1985). Competitive Advantage. *Competitive Advantage: Creating and Sustaining Superior Performance*. http://doi.org/10.1182/blood-2005-11-4354

Porter, M. E. (2008). Value-Based Health Care Delivery, *248*(4), 503–509. http://doi.org/10.1097/SLA.0b013e31818a43af

Porter, M. E. (2009). A Strategy for Health Care Reform — Toward a Value-Based System. *The NEW ENGLAND JOURNAL of MEDICINE*, *363*(1), 1–3. http://doi.org/10.1056/NEJMp1002530

Porter, M. E. (2010). What Is Value in Health Care? *The New England Journal of Medicine*, 2477–2481. http://doi.org/10.1056/NEJMp1415160

Porter, M. E., & Millar, V. E. (1985). How information gives you competitive advantage. *Harvard Business Review*.

Porter, M. E., Pabo, E. A., & Lee, T. H. (2013). Redesigning Primary Care: A Strategic Vision To Improve Value By Organizing Around Patients' Needs. *Health Affairs*, *32*(3), 516–525. http://doi.org/10.1377/hlthaff.2012.0961

Porter, M. E., & Teisberg, E. O. (2006). Redefining health care: Creating value-based competition on results. *Atlantic Economic Journal*, *35*(4), 491–501. http://doi.org/10.1007/s11293-007-9091-9

Prodel, M., Augusto, V., Xie, X., Jouaneton, B., & Lamarsalle, L. (2015). Discovery of patient pathways from a national hospital database using Process Mining and Integer Linear Programming, 1409–1414.

Rebuge, A., & Ferreira, D. R. (2012). Business process analysis in healthcare environments: A methodology based on process mining. *Information Systems*, *37*(2), 99–116. http://doi.org/10.1016/j.is.2011.01.003

Rozinat, A., & Aalst, W. M. P. van der. (2005). Conformance Testing: Measuring the Fit and Appropriateness of Event Logs and Process Models. *BPM 2005 International Workshops, BPI, BPD, ENEI, BPRM, WSCOBPM, BPS*, 163–176. http://doi.org/10.1007/11678564_19

Rubin, H. R., Pronovost, P., & Diette, G. B. (2001). From a process of care to a measure: the development and testing of a quality indicator. *International Journal for Quality in Health Care : Journal of the International Society for Quality in Health Care / ISQua*, *13*(6), 489–496. http://doi.org/10.1093/intqhc/13.6.489

Shewhart, W. A. (1980). *Economic Control of Quality of Manufactured Product*. American Society for Quality Control. Retrieved from https://books.google.nl/books?hl=en&lr=&id=XBeoAgAAQBAJ&oi=fnd&pg=PR1&dq=Economic+Control+of+Quality+of+Manufactured+Product&ots=PliZYLVoWA&sig=YiA08ZkrPPaDqDhQYN-

meSydU3E&redir_esc=y#v=onepage&q=Economic Control of Quality of Manufactured Product&f=false

Smit, M. (2016). Wat zijn de vijftien grootste ziekenhuizen van Nederland? Retrieved from http://www.vektis.nl/index.php/publicaties/onderzoeksresultaten/37-wat-zijn-de-vijftien-grootste-ziekenhuizen-in-nederland

Strausbert, J., Bilir, H., Waydhas, C., & Ruchholtz, S. (2003). Guideline validation in multiple trauma care through business process modeling. *International Journal of Medical Informatics*, *70*(2–3), 301–307.

The Dutch Council for Health and Care. (2013). *Guarantees for quality of care*.

van Deen, W. K., Esrailian, E., & Hommes, D. W. (2015). Value-Based Health Care for Inflammatory Bowel Diseases. *Journal of Crohn's and Colitis*, 421–427. http://doi.org/10.1093/ecco-jcc/jjv036

van den Berg, M., de Boer, D., Gijsen, R., Heijink, R., Limburg, L., & Zwakhals, S. (2014). *Zorgbalans 2014*. *Rijksinstituut voor Volksgezondheid en Milieu*.

van der Aalst, W. M. P., de Beer, H. T., & van Dongen, B. F. (2005). Process mining and verification of properties: An approach based on temporal logic. *Springer*, *3760*, 130–147.

van der Aalst, W. M. P., Reijers, H. A., Weijters, A. J. M. M., van Dongen, B. F., Alves de Medeiros, A. K., Song, M., & Verbeek, H. M. W. (2007). Business process mining: An industrial application. *Information Systems*, *32*(5), 713–732. http://doi.org/10.1016/j.is.2006.05.003

Van Der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, *16*(9), 1128–1142. http://doi.org/10.1109/TKDE.2004.47

volksgezondheidenzorg.info. (2015). Feiten en cijfers over volksgezondheid en zorg. Retrieved from https://www.volksgezondheidenzorg.info/

11 GLOSSARY

Diagnose = determining a certain state-based on data.

Patient outcome = the data that is generated after each performed process.

Health outcome = Health outcome is the outcome that is stated for a certain medical condition. For example, obese patients have losing weight as a health outcome, cancer patients have surviving cancer as a health outcome, etc.

Treatment = Consists of a number of care activities. This package of care activities is linked to, a so called, Diagnosis Treatment Combination (in Dutch Diagnose-Behandel Combinatie, DBC).

DBC = the payment of hospital care based on the Diagnosis Treatment Combination.

Care process = a process consists of different activities that are carried out in a predefined order.

Care activity = what needs to be carried out to treat a patient with a certain diagnose.

Caregiver = Defined broadly as any person within the healthcare setting whose work touches a patient's or family's care experience, including the doctors, nurses, therapists, technicians, dietitians, appointment schedulers, parking attendants, janitors, and even hospital leaders, purchase and supply chain employees, and financial representatives who the patient and family might never see (DiGioia, 2012).

Outcome indicator = determining mortality or complications as a result of treatment, but also the achieved health gain and its sustainability (Kuenen et al., 2011).

Process model = a view of a flow or series of activities.

Process maps = the set of process models.

Process = a flow or series of activities.

Activity = A task (can be pre-defined) to be performed.

"An activity is an action that is a semantical unit at some level. In addition, each activity can be thought of as a function that modifies the state of the process" (Agrawal, Gunopulos & Leymann, 1998).

Events = An event in a case.

12 APPENDICES

12.1 COMPUTE THE TOTAL COSTS OVER EACH PATIENT'S CYCLE OF CARE

The formula looks like the following.

1. Calculation of resource capacity costs:

Total monthly expenses / (Total monthly days * Total available hours per day).

2. Calculation of the total cost capacity per patient per care cycle:

amount of time spent (per hospital personnel) * resource capacity costs.

Below are examples with aspects with which the cost of processes can be calculated.

Example of cost per expert per year and month						
	Annual compensation (with secondary working conditions)	Supervision fees	Occupation of spaces	Technology and Support	**Total annual cost**	**Total monthly cost**
Nurse	65000	9000	10800	2560	**87360**	**7280**

Table 5: Example of cost per expert per year per month

Example of time (in days) per expert per year and per month						
	Start days	Weekend days	Leave days	Holidays	Sick days	**Available days per year**	**Available days per month**
Nurse	365	104	20	12	5	**224**	**18.7**

Table 6: Example of time (in days) per expert per year per month

Example of time (in hours) per expert per day				
	Start hours	Scheduled breaks (hours)	Meetings, training and education	**Total available hours per day**
Nurse	7.5	0.5	1.0	**6**

Table 7: Example of time (in hours) per expert per day

Calculating the capacity cost rate per

- Calculation of resource capacity costs: 7280 / (18.7 * 6) = 64.88
- Calculation of the total cost capacity per patient per care cycle: 0.4 (hours spent with nurse by patient X) * 64.88 = 25.95. This is the cost for one patient in a care cycle who has visited a nurse.

12.2 THEORETICAL MODEL OF CASE MAMMACARCINOOM

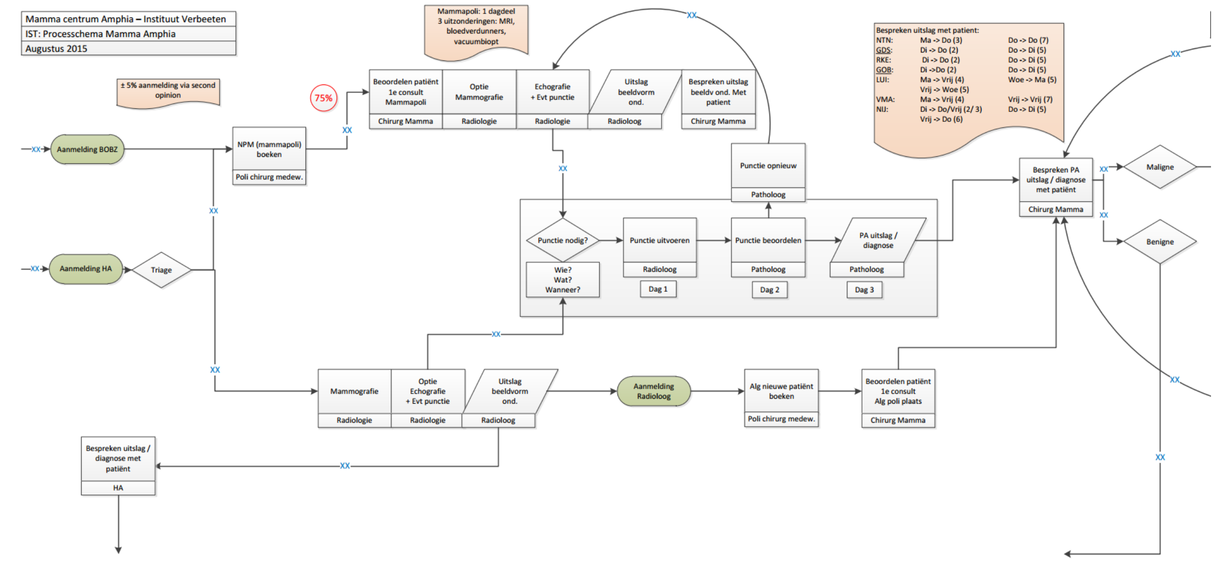


Figure 47: Theoretical model of Case Breast Cancer (part 1)

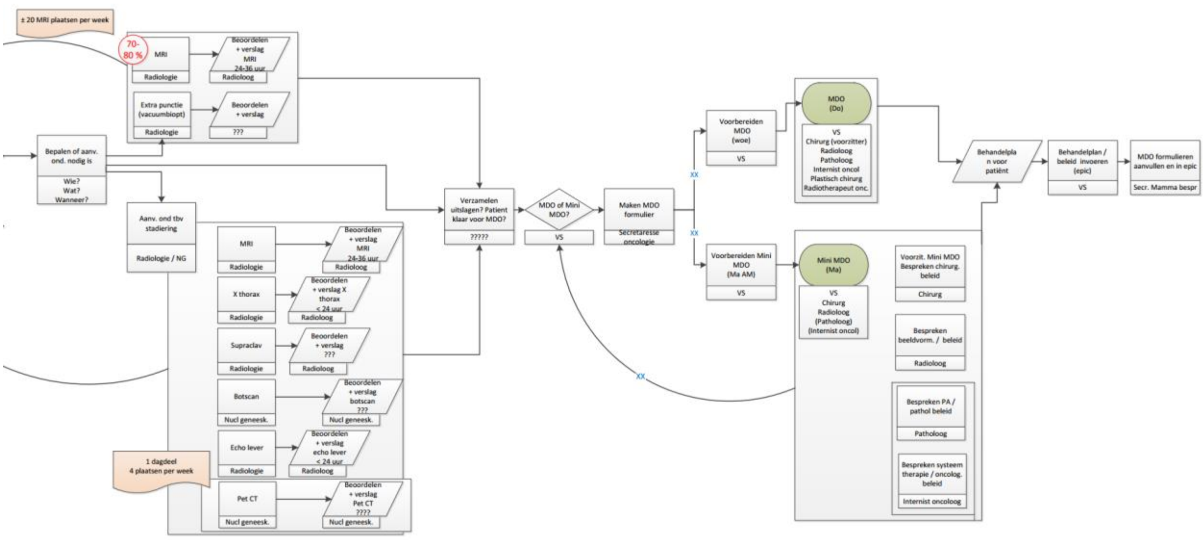


Figure 48: Theoretical model of Case Breast Cancer (part 2)