Utrecht University

Department of Information and Computing Science

Master Thesis for the Master's degree of Business Informatics

A Comparative Study of Process Pattern Discovery Frameworks in a Medical Setting

First supervisor:
Dr. ir. X. Lu

Author:
Sabrina Türker

Second supervisor:
Dr. G.C. van de Weerd

In cooperation with:
Maxima Medisch Centrum

August 14, 2024

Abstract

In the last decade several Process Pattern Discovery methods have been developed, such as Sequential Pattern Mining and Significant Pattern Mining. These methods aim to extract patterns from event logs that are of interest to the user. This is done by employing either a single or multiple interest dimensions. Each Process Pattern Discovery methods have their own algorithms and software to extract these process patterns from a event log. By comparing these methods with each other in a medical setting, their strengths and weaknesses can be better understood, allowing for the identification of the most effective method for medical event logs. Hence, a comparative study will be performed. Therefore, this thesis proposal explores the application of three Process Pattern Discovery methods: the IMPResseD framework, Episode Discovery, and Local Process Model. These methods will be used to analyze Novasure surgeries through process pattern discovery, each offering its own approach to uncovering patterns. The IMPResseD framework employs a multi-interest function, whereas Episode Discovery focuses on the frequency of occurring patterns within the event log, and the Local Process Model utilizes five quality criteria. To compare these three methods, an evaluation plan will be implemented. This plan will assess validity and quality through both qualitative and quantitative analyses. Qualitative analysis involves conducting an in-depth interview with a domain expert, while quantitative analysis will utilize cross-validation on the discovered patterns to assess the predictive capabilities of the three methods. Ultimately, this research aims to provide valuable insights into the effectiveness of these methods in discovering patterns that result in needing an intervention after the Novasure surgery is performed.

Acknowledgements

I would like to express my heartfelt gratitude to the individuals who have significantly contributed to the successful completion of my master's thesis in Business Informatics.

First and foremost, I want to thank Suhwan Lee for creating and providing the merged event log about the Novasure. You saved me a lot of time and effort, so thank you very much.

I am deeply grateful to Mozhgan Vazifehdoostirani for their support with the IMPresseD framework. Your patience in explaining the software and its functionalities, as well as answering my numerous questions, was crucial in continuing my work for my master thesis.

I would also like to thank Inge van de Weerd for serving as my second supervisor and providing valuable tips and feedback throughout the process of writing my thesis.

A special thank you goes to Jaklien Leemans, who generously found time in their busy schedule to be interviewed. Your insights into process patterns and the literature you provided on the Novasure surgery were incredibly enlightening and added depth to my research.

Most importantly, I owe a profound debt of gratitude to my supervisor, Xixi Lu. Your unwavering support, weekly guidance, and willingness to address all my questions, doubts, and problems were the backbone of this thesis. Your mentorship has been invaluable, and I am deeply appreciative of your commitment to finishing my master thesis.

Lastly, I want to express my heartfelt thanks to Quinesha van Burgh for supporting me throughout this journey. You were there for me through the highs and lows, and I am deeply grateful for your loving support.

Thank you all for your contributions; without your support, this thesis would not have been possible.

Contents

List of Figures

List of Tables

List of Acronyms

AUB	**A**bnormal **U**terine **B**leeding
EP	**E**pisode **D**iscovery
GUI	**G**raphical **U**ser **I**nterface
IMP	**IMP**resseD Framework
IMPresseD	**I**nteractive **M**ulti-Interest **P**rocess Pattern **D**iscovery
LPM	**L**ocal **P**rocess **M**odel
MMC	**M**axima **M**edical **C**enter
PPD	**P**rocess **P**attern **D**iscovery
M	**M**edical
NM	**N**on-**M**edical

1. Introduction

This chapter emphasizes the role of process mining in healthcare, specifically in enhancing clinical processes' performance and outcomes. It focuses on Process Pattern Discovery (PPD), which extracts process patterns from event logs. The research thesis targets the analysis of a dataset from the Maxima Medical Center (MMC) in Eindhoven, focusing on patients undergoing an intervention, namely the Novasure surgery, for Abnormal Uterine Bleeding (AUB). Despite the surgery's aim to reduce menstrual bleeding by removing the uterine lining, complications persist. PPD methods are utilized to identify process patterns potentially leading to these complications.

Outlined in the thesis are four PPD methods: Episode discovery, Local process model, IMPresseD framework, and Signature discovery framework. These PPD methods are chosen based on their strategy of discovering process patterns, such as frequency or feature selection. The objective is to evaluate and compare the accuracy and alignment of these methods' outcomes and pinpoint potential factors contributing to post-surgery complications. Accuracy will be acquired via cross-validation, while alignment refers how much the order and combination of events within a process pattern is in line with the expectation of the medical expert.

1.1 Background

Process mining consist of a set of techniques for evaluating the efficiency and compliance of business processes by examining event logs, which is a collection of actions or occurrences recorded by a software system, generated during their execution. These techniques variate from automated process discovery, where a business process model is produced based on an event log, to conformance checking, where the input is a process model and an event log and the output are the differences between the model and the event log [1]. Process mining has a crucial contribution to understanding the bottlenecks and performance of a process by extracting knowledge from event logs. This enables organizations to identify issues and generate solutions for them. It also aids in identifying outliers and, more importantly, understanding the reasons behind their deviation [2].

A field where process mining is used frequently is healthcare. Due to the complex and life-threatening nature of healthcare, process mining can have

a tremendous impact on the performance and outcomes. The application of process mining in healthcare is on various levels, including the analysis of organizational processes such as the scheduling of doctors and nurses' shifts [3]. It also delves into clinical processes, focusing on the medical treatments that patients undergo throughout their stay in the hospital [4].

Within process mining several sub fields exist, one of them is process pattern discovery (PPD) which focuses on discovering process patterns from event logs. It is well known that some processes, such as treatment trajectory in healthcare, exhibit a chaotic structure [5]. In addition, a vast amount of data is generated by hospital staff and the systems they use. However, there is currently no systematic approach to handling this large volume of data using process mining, except in case studies within a research context [6]. This can result in end-to-end models that are hard to read for the hospital staff, because these models can include a lot of events while process pattern consist of only a few events. Thus, in these cases, it is more practical to look at process patterns within a process, instead of looking at the end-to-end processes. Furthermore, the methods to identify these process patterns can be adjusted to the interest of the user, such as the Interactive Multi-Interest Process Pattern Discovery (IMPresseD) framework [7].

One example where PPD is applied in a clinical setting is described in the paper by Huang, Lu and Duan [8]. The paper focuses on uncovering hidden or latent process patterns within clinical processes by using event logs. The goal is to identify treatment patterns that may not be immediately apparent. This research facilitated the discovery of hidden treatment patterns and the grouping of similar process patterns to recommend comparable treatments to patients.

Several methods exist to extract and analyse process patterns within a process model. In this thesis the focus will lie on four methods, which are Episode discovery, Local process model, IMPresseD framework, and Signature discovery framework. The details of these methods will be described in chapter 3.

1.2 Novasure Surgery

This research will be performed in collaboration with the Maxima Medical Center (MMC) in Eindhoven. The MMC provided a dataset of patients with several characteristics and their outcomes, namely if an intervention,

the Novasure surgery, was performed or not. The Novasure is a surgery for people who suffer from heavy menstrual bleeding, also called Abnormal Uterine Bleeding (AUB). This treatment reduces or stops menstrual bleeding by permanently removing the uterine lining. This is done by rapid release of radio frequency energy which is performed by a device called the Novasure, hence the name of the surgery. Due to the removing of the uterine lining, the Novasure surgery is only performed on people who do not have a desire to have children [9]. More information on the dataset will be discussed in chapter 4.2.

AUB is observed in 11% to 20% of the general population, with a higher occurrence in individuals of increasing age or during their perimenopausal phase which refers to the transitional period leading up to menopause. A general practitioner identifies menstrual bleeding as AUB when it deviates from normal standards across four indicators: the frequency and regularity of the menstrual cycle, the length of the period, the duration of flow in days, and the volume of monthly blood loss measured in milliliters [10]. AUB has negative effects on health, leading to anemia, a reduced number of red blood cells, and significantly influences the quality of life for the affected people. Additionally, it carries economic consequences for both individuals such as the cost of sanitary pads, and society including expenses related to sick leave and surgical costs, for example the Novasure surgery [11].

An intervention such as the Novasure surgery is an outcome for patients who suffer from AUB. However, complications and complaints from patients still exist after the Novasure surgery is performed. Some of these complications involve persistent bleeding even a year after undergoing Novasure surgery, leading to re-interventions such as follow-up surgery. In certain cases, a hysterectomy might be required [12]. Therefore, it is of utmost importance to address the issue and discover which factors contribute to complications after an intervention.

PPD can help discover these potential factors by analysing process patterns within a patients event log to see which process patterns might lead up to complications or complaints after the Novasure surgery is performed.

1.3 Outline Thesis

This thesis is structured in the following way, Chapter 1 discusses the introduction of the thesis, specifically the background of process pattern discov-

ery and the process pattern detection methods, Novasure dataset and the outline of the thesis. Chapter 2 discusses the problem statement which contains the research questions and the motivation for this research. Chapter 3 dives deeper into the four methods of process pattern discovery and their differences. Chapter 4 will discuss the research method which consist of a description of the dataset and its preprocessing, evaluation setup and the limitations of this research method. Chapter 5 consist of the results of the evaluation setup. In Chapter 6 the discussion will take place of the research method. Lastly, in Chapter 7 the conclusion will be discussed.

2. Problem Statement

In this chapter, the problem statement will be addressed, and appropriate research questions will be presented to resolve this issue. The primary focus of this research is to bridge the scientific gap in information science, specifically through a comparative analysis of different PPD methods within a medical setting. Therefore, a main research question is formulated as well as several sub-questions to answer the main research question. These sub-questions focus on the accuracy and alignment of the discovered process patterns, as well as the contributions of these process patterns to the intervention complications.

2.1 Motivation

The outcome of an intervention is not always favorable for patients, and re-interventions might be necessary for them. It is essential to understand the underlying factors leading to re-interventions.

Additionally, there is a gap in the scientific community concerning comparative studies of various PPD methods, which could reveal their relative strengths and weaknesses. A comprehensive comparative study is therefore beneficial for evaluating the effectiveness of different PPD methods in a medical setting. This research aims to address this gap by using the Novasure dataset to compare these methods.

By doing so, the study will contribute valuable insights to the field of process pattern mining in healthcare, providing practical recommendations for medical experts and researchers on the most effective PPD methods for analyzing healthcare processes. Furthermore, the findings could pave the way for future research into the development of more advanced PPD methods specifically tailored to medical applications.

2.2 Research Questions

To address the aforementioned issues, the following main research question is formulated:

How do various process pattern detection methods perform on Novasure data set,

and what insights can be gained through this comparison?

In order to answer the main research question, a set of sub-questions has been formulated:

1. Which of the discovered process patterns relate and contribute to the cases that require re-interventions?
2. How do various process pattern discovery methods compare in their performance for discovering process patterns?
   (a) To what extent are the discovered process patterns aligned with the medical expert's expectations?
   (b) Performance-wise, what are the differences in the discovered process patterns across various process pattern discovery methods, in terms of accuracy?

The strategy for answering the main research question and its sub-questions consists of performing a qualitative and quantitative analyses on the discovered process patterns. Sub-question two (a) will be addressed through qualitative analysis, which consists of selecting three discovered process patterns of each PPD method. These process patterns will then be used during an interview with a medical expert to discuss the extent to which the discovered process patterns align with their expectations. Sub-question two (b) will be answered via quantitative analysis. This consists of analyzing the top three process patterns of each PPD method via cross-validation, which will result in accuracy measurements. These measurements will then be used to assess the quality of the process patterns in predicting re-intervention or not. Further details on these approaches will be presented in Chapter 4.1.

3. Related Work: Process Pattern Discovery Methods

This research will focus on four PPD methods, which are Episode Discovery, Local Process Model, IMPresseD Framework, and Signature Discovery Framework. All these four frameworks focus on discovering patterns within an event log. The selection of these frameworks, the frameworks them self and their differences will be discussed below.

3.1 Selection of PPD Methods

While process mining brings a lot of advantages, several challenges still exist. One of these challenges are inaccurate process models that are not aligned with domain experts. For example, in a medical setting, doctors or nurses may record events in a system in a certain order, but from the patient's perspective, the order may be different. Therefore, it is crucial to examine the process from the patient's viewpoint [6]. Another challenge is the inability to extract insights from them due to the process representation, because the process model has too many connections making it hard to read [13].

Over the last decade, different PPD methods have been developed to handle these limitations, such as the frequent sequence pattern. The results in the paper show a simpler and more insightful process map by only showing crucial patterns [14]. Each PPD method has its own strategy to extract certain patterns, ranging from focusing on frequency [5] to action-response-effect [13].

In this research, the focus will lie on four PPD methods, which are chosen based on their strategy of discovering process patterns. The IMPresseD Framework is unique in itself because it combines multiple interest functions to interactively discover patterns, while other methods focus on one single interest function [7]. Therefore, this method will be included in this research. Episode Discovery is selected because it focuses on the frequency of patterns within an event log [5], which is also an interest function in the IMPresseD Framework. Therefore, it can result in an interesting comparison to see if they output the same type of results. The Local Process Model discovery process patterns that occur frequently within a process pattern

but it also looks at behavioral patterns, meaning it includes sequential composition, concurrency, choice, and loop, just like process mining but focused on process pattern instead of start-to-end process models. Because of this, it can result in richer patterns, which would be interesting in this comparison. Furthermore, it also focuses on frequency, which is interesting to compare with the aforementioned PPD methods which also use this [15]. The Signature Discovery Framework focuses on finding patterns that can differentiate between desirable and undesirable behavior. This is done by performing feature selection and class labeling, namely faulty or normal behavior. This PPD method does not use frequency at all and uses data science techniques such as decision tree learning and association rule mining to acquire the process patterns [16]. This can result in discovering interesting process patterns; therefore, it is included in this research.

Several more PPD methods exist but will not be used in this research due to time limitations and the specific purposes of the PPD methods. An example of a PPD method that is excluded, is the action-response-effect method to discover patterns. The focus of this method lies on the representation of the process model to ensure that domain experts can extract crucial insights from it. This is achieved by using statistics to find process patterns that have an action, response, and effects characteristic. However, in this comparative research, the focus lies on extracting insightful patterns for comparison, rather than on the creation of understandable process models for the domain experts [13].

Another example of a PPD method that will not be used is the Guided Process Discovery method. This method focuses on translating low-level events to high-level events by grouping the low-level events into an activity pattern with the help of domain experts. In this case, the definition of a process pattern is used differently, and the focus lies on acquiring insights for domain experts by grouping low-level events into high-level events to make them more understandable. However, the focus of the comparison study performed in this research lies in finding the differences among types of PPD methods that extract process patterns from an event log, rather than making them more explainable to domain experts [17]. Therefore this method will not be used in this research.

3.2 Episode Discovery

The Episode Discovery is described in the following paper [5], a method positioned at the intersection of process mining and pattern mining because the paper uses ideas from episode mining. It focuses on discovering frequent episodes within a process by using partial orders. Moreover, it discovers episode rules to forecast behavior and identifies correlated behaviors in the process. According to the paper the definition of an episode is a partially ordered collection of events, in other words it looks at a local behaviour of a process within an event log. This is done by using a combination of various algorithms at each step, these steps and algorithms will be discussed below:

1. **Frequent Episode Discovery**: this algorithm is used to discover a set of frequent episodes in the event log.
2. **Episode Candidate Generation**: the purpose of this algorithm is to use the set of frequent discovered episodes as input and output candidates episodes by merging overlapping episodes with each other that share all the nodes and several edges.
3. **Frequent Episode Recognition**: the goal of this algorithm is to check whether the episode candidates are frequent by using a frequency threshold. The episode candidates that are not frequency will be filtered out.
4. **Time Complexity Analysis**: for the algorithms in the Frequent Episode Discovery and Frequent Episode Recognition the worst case time complexity is calculated, however the Episode Discovery algorithm is quite fast in practice. In general this is not part of discovering episodes within an event log, but an in between step to check the time it takes for these algorithms to perform.
5. **Pruning**: the purpose of this algorithm is to filter out any uninteresting episodes, which also minimize the computation time. This is done by several pruning techniques, which are activity pruning, trace distance pruning, pruning based on the antisymmetry of $\leq$ and pruning based on the eventually-follows relation.
6. **Episode Rule Discovery**: the purpose of this algorithm is to find a set of valid episodes rules by using a discovery tree. In this discovery tree the found episodes are presented and a parent-child relationship can be derived from it, which results in the episode rules.

3.3 Local Process Model

The Local Process Model is described in the paper [15]. This PPD method discovers frequent behavioural patterns in processes. The focus of the Local Process Model is smaller and frequent patterns of around three to five activities which are presented in a process tree. This is done by using five quality criteria to present the representation of a pattern within an event log. This approach consist of four steps, where the last three steps can be repeated. These steps are discussed below:

1. **Generation**: the purpose of this step is to generate candidate Local Process Models which are represented in a process tree. These trees consist of leafs where each leaf is one activity. The first round starts out with one leaf, thus one activity. This results in a set of Local Process Models.
2. **Evaluation**: the Local Process Models are evaluated based on five quality criteria which are:
   (a) Support: the number of patterns in an event log that are part of a Local Process Model. So if a Local Process Model has a higher support than any other Local Process Model, it means that is appears frequently in the event log and thus represents it better.
   (b) Confidence: is the ratio of an event type that fit a specific pattern in an event log to the total number of events in an event log.
   (c) Language fit: the ratio of observed behaviour in an event log that is allowed by a Local Process Model. If a Local Process Model allow for much more behaviour, then chances are high it overgeneralizes. Which results in not describing the behaviour in the event log well.
   (d) Determinism: this relates to the certainty of a Local Process Model. A Local Process Model with a high degree of certainty, contains a clear path. For example when a Local Process Model contain a lot of XOR gateways, more possible options are available which creates a higher uncertainty.
   (e) Coverage: the ratio of the amount of events in the log after projecting all the events to a Local Process Model to the total number of events in an event log.

   Furthermore two quality dimensions are used which are:

(a) Local Process Model selection & weighted average: even though there are five quality metrics some metrics might be more valuable than others depending on the analysis goals. Therefore it might be useful to use a weighted average of the five quality metric for selecting a Local Process Model.

(b) Monotonicity proprieties & pruning: when a Local Process Model contains low values for several quality metrics, it might be a good idea to prune these models away because expanding them still leads in low quality metric values due to monotonicity.

3. **Selection**: based on the results of the quality criteria, a selection is made of the Local Process Models.

4. **Expansion**: the selected Local Process Models are expanded, by adding an operator node and another activity. After this step, the evaluation, selection and expansion steps can be performed again on the expanded Local Process Models. This procedure stops until no process tree meets the quality criteria anymore.

3.4 IMPresseD Framework

The IMPresseD Framework is described in the paper [7]. It uses a multi-interest-driven framework to discover patterns that are of interest to the user. This is performed by combining different techniques and defining multiple interest functions by the user. In the paper the framework was applied on a hospital dataset of cancer patients, similar to the hospital setting of the Novasure dataset used in this research. The IMPresseD Framework consist of the following steps:

1. **Converting traces**: the first step is to convert the traces in the event log into partially ordered traces. This is done by using a conversion oracle which is derived from expert knowledge or data analysis.

2. **Defining interest functions**: the purpose of this step is to define interest functions which are in accordance with what the user find interesting with regards to patterns. These patterns are aligned with the analysis goal of the user. The following interest functions are presented in the paper:

   (a) Frequency interest: which calculates the frequency of a pattern in an event log.

   (b) Outcome interest: which measures the effect of each pattern on the process outcome.

   (c) Case distance interest: which is used to mitigate the impact of confounding variables.

3. **Extracting patterns**: the goal of this step is to extract patterns of length one which are, for this length, individual activities.
4. **Measuring the interest functions**: in this step the interest functions are calculated for the extracted patterns.
5. **Returning patterns**: the objective of this step is to return the set of patterns that are optimal according the the calculations derived from the interest functions. The optimal patterns are found by using the Pareto Front of all the patterns.
6. **Checking patterns**: the purpose of this step is to check the set of patterns with the user, which is done by using visualizations. The user can then decide if it is satisfied with the current set of patterns or not. If the user is satisfied, the procedure will end here. If not, the user will select patterns from the set of patterns to be extended upon, which results in the following step.
7. **Extending patterns**: the last step of this method consist of extending the selected patterns and return to step 4, measuring the interest function again but with the extension.

3.5 Signature Discovery Framework

The Signature Discovery Framework is described in the paper [16] and is used to discover signatures, which are patterns that discriminate between different classes of behaviour. In other words, these patterns are a key component for prediction models or explaining certain type of behaviour. The framework consist of several steps, which are discussed below:

1. **Event log**: the starting point is an event log comprising ordered events, where each event corresponds to a well-defined step in a process related to a specific case. This log captures diverse attributes of events such as timestamps and resources.
2. **Class labeling**: all instances in the event log should be assigned a class label to discover class-specific patterns and have distinctions between classes. When dealing with unlabeled cases in event logs, cluster-

ing and classification techniques from machine learning can be used. Methods such as k-nearest neighbor or one-class Support Vector Machines (SVM) are applied for automatic or semi-automatic labeling.

3. **Feature extraction and Selection**: the focus is on deriving features from the event log. This involves transforming each event log instance into a vector space where the vector elements represent the values of selected features.
4. **Discover patterns**: the objective of this step is to identify patterns across features strongly correlated with the class label such as normal or faulty. This steps employs standard data mining techniques, specifically decision tree learning and association rule mining.
5. **Evaluation**: metrics in data mining, including true positives, false positives, true negatives, and false negatives, are employed. Derived metrics such as accuracy, sensitivity, specificity, precision, and F1-score are then utilized to assess the quality of the discovered signatures. To address potential overfitting issues and ensure model generalizability, cross-validation techniques are applied during the learning phase.
6. **Reporting and Visualization**: the framework concludes by presenting findings and visualizing results. Automated reports, showcasing signature patterns along with their performance metrics, are generated.

3.6 Differences

By discussing each method individually, as done previously, the differences between them can be clarified in this section. Almost all four PPD methods have their own naming for process patterns along with their respective definitions. According to the IMPresseD Framework, process patterns consist of a set of activities with an ordering relation, allowing for additional annotated data [7]. In contrast, the Episodes Discovery uses the name episodes to describe a partially ordered collection of activities from an event log [5]. The Local Process Model refers to them as Local Process Models, focusing on a subset of process activities within a process model and describing specific behavioral patterns that are frequently found within an event log [15]. Lastly, the Signature Discovery Framework labels them as signature patterns, capable of distinguishing between various classes of behaviour, such as normal or faulty behaviour, based on their features [16]. Essentially, these are different names for process patterns, but the way they are discovered

varies due to the algorithms and strategies used for each respective PPD method.

Both the IMPresseD Framework and Episode Discovery mention the term "order" in their definitions, whereas the Local Process Model and Signature Discovery Framework do not. This difference is logical considering that the Signature Discovery Framework employs machine learning techniques to discover process patters, and the Local Process Model utilizes quality criteria in combination with expanding the local process model by adding an operator node or another activity after each expansion. Therefore, their focus is less on order, though it remains important.

Another observation is that all methods, except the Signature Discovery Framework, discuss a set, subset, or part of activities from an event log. This emphasizes that they deal with a small portion of activities from an event log rather than a complete trace, such as start-to-end process models. For the Signature Discovery Framework, this is less the case because its focus lies on uncovering common patterns that result in faulty or normal behavior.

Another way to look at the differences of the PPD methods, is their focus. Each method has its own focus when it comes to discovering process patterns. The IMPresseD Framework employs interest functions to discover process patterns, while Episode Discovery utilizes frequency to uncover process patterns, and the Local Process Model employs statistics with five quality criteria to discover process patterns. Lastly, the Signature Discovery Framework utilizes class labeling, feature selection, and machine learning to uncover process patterns.

There are more differences to be found within the four PPD methods. Therefore, five criteria have been chosen to delve deeper into the differences among these methods. These criteria and their description will be discussed below:

1. **Use of frequency**: the use of frequency for discovering process patterns is important because if frequency is used within the PPD method, it could potentially omit important patterns that are less frequent within an event log.
2. **Use of domain experts**: the inclusion of domain experts when discovering process patterns is crucial because domain experts possess the knowledge of how a process should be, thus enabling them to determine the value of the discovered process patterns. Their input could

result in more insightful process patterns. In this research setup, the term medical expert will be used instead of domain expert due the the medical setting of this research.

3. **Adaption of search**: the option to adapt the discovery of process patterns by adding or changing a function or measurement could be useful for domain experts to tailor the search to their analysis goals. Therefore, it is named adaptation because the domain expert is able to adapt the function or measurement for the iteration of the process pattern discovery.
4. **Output of case ID**: The output of the case ID with the discovered process patterns is a valuable function for domain experts, making the PPD methods more user-friendly for them.
5. **Narrowing down technique**: another criterion is the method used to narrow down the discovered patterns, such as pruning or starting small and increasing the number of patterns each round by using filters or thresholds. This could lead to slightly different process patterns depending on who performs it or how it is executed.
6. **Use of iterations**: lastly, the potential for iterations with the discovered patterns from the initial round to refine the selection is essential as it can enhance the current set of process patterns, leading to qualitatively better patterns.

See Table 3.1 for an overview of the aforementioned criteria and how they compare with the four PPD methods.

Lastly, their strategies and the steps for finding process patterns differ. The IMPresseD Framework and the Local Process Model both incorporate an extension step where they add an operator node or activity and iterate through the steps again to assess the quality of the patterns. In terms of filtering out patterns, for the IMPresseD Framework, this is typically done by a domain experts who check the discovered process patterns. For Episode Discovery, it involves pruning the discovered process patterns. For both the Signature Discovery Framework and the Local process, it is similar to their evaluation step. In the case of the Local Process Model, the evaluation of the process patterns is based on the five quality criteria, after that the operator nod or activity will be added. The Signature Discovery Framework evaluates the patterns using metrics in data mining to assess their quality after the discovery has been performed. For Episode Discovery, evaluation is not as explicit; it involves using different algorithms such as frequency and ver-

Table 3.1: The chosen criteria for comparing the PPD methods

PPD method	**Use of frequency**	**Use of domain experts**	**Adaptation of search**	**Output of case ID**	**Narrow down technique**	**Use of iterations**
IMPresseD Framework	Yes	Yes	Yes	Yes	Filtering down by domain expert	Yes
Episode Discovery	Yes	No	No	No	Pruning	No
Local Process Model	Yes*	No	Yes	No	Refining***	Yes
Signature Discovery Framework	No	Yes**	No	No	Not performed	No

* It is named support, which is one of the five quality criteria of the Local Process Model.

** Only involved in the feature selection.

*** Ultimately, the Local Process Model adds an operator node or activity after each iteration and then reevaluates against the five quality criteria until no patterns fit the quality criteria. Therefore refining is more fitting in this context.

ifying if they are indeed frequent by applying a threshold. After that the pruning step takes place.

3.7 Discussion

All four methods share the common goal of uncovering patterns within an event log, with their own respective methods. However, differences can be found within these methods.

The IMPresseD Framework uses multiple interest functions to discover patterns. These interest functions are defined to align with the user's analysis goal, potentially resulting in interesting pattern discovery during the research. This is why the IMPresseD Framework will be used during this comparison.

Episode Discovery relies on frequency to find patterns. Therefore, patterns that occur more often within an event log will be the outcome. This is why Episode Discovery will be used in this research to assess if frequency is a suitable metric for importance.

The Local Process Model employs several quality metrics to identify representative patterns of an event log. This might result in different pattern outcomes compared to the aforementioned methods, potentially leading to interesting pattern discovery during the research.

The Signature Discovery Framework uses class labeling and data mining techniques to find patterns that highly correlate with the labeled behaviour. However, the Signature Discovery Framework focuses on identifying signatures that distinguish various classes of behavior, such as faulty and normal, in contrast to other methods that extract patterns based on frequency, interest functions, or quality metrics. Consequently, the Signature Discovery Framework does not use a singular dimension to identify patterns. Instead, it utilizes selected features to identify patterns which highly correlate with the labeled behavior. Therefore leaving out features that might be crucial to the outcome of an intervention. Additionally, the Novasure dataset contains a tremendous amount of features. Thus it is not reasonable to select a few features by the user, because it might lead to biased results during the pattern discovery. Therefore no interesting and new patterns will be discovered. As a result of this, the Signature Discovery Framework will not be used in the continuation of this research.

3.8 Summary

Overall, the four PPD methods are discussed elaborately and three of them will be used in the continuation of this research. There are differences in the naming of process patterns, as well as in the strategies and algorithms used to discover these patterns across different PPD methods. The impact of these differences on the outcomes of process pattern discovery is still unclear. Therefore, a comparative study is beneficial to evaluate the strengths and weaknesses of various PPD methods within a medical setting. Such a study would provide valuable insights into the field of process mining in healthcare. In addition, it would close the gap within the scientific community with regards to the various PPD methods and their effectiveness on, specifically, a medical dataset.

To date, no comparative study exists that focuses on the variety of PPD methods. However, research has been conducted comparing different process mining tools available on the market [18]. Additionally, studies have compared different processes rather than focusing on a single process [19]. Nonetheless, a comparative study on PPD methods is still missing to this date in the scientific community.

4. Research Method

This chapter discusses the research methods for comparing the three different PPD methods. Firstly, information about the dataset that will be used for the evaluation plan is provided. Additionally, an evaluation plan comprising a systematic selection on the discovered process patterns, a qualitative and quantitative analyses, is developed for this research. The specifics of this dataset, evaluation plan, and its analyses will be discussed in the following sections below.

4.1 Evaluation Overview

The following section details the steps involved for the comparative evaluation of the three PPD methods. This evaluation setup consists of three parts.

The evaluation setup first involves a systematic selection of the discovered process patterns. This is due to the variety of types of events within the dataset. The outcome of the systematic selection will then be used for both the qualitative and quantitative analyses. Afterwards, a qualitative and quantitative analyses will take place to assess the quality and validity of the results from the three PPD methods. Qualitative analysis examines the nature and characteristics of the findings, while quantitative analysis focuses on numerical data to assess reliability and significance. This combined approach provides a comprehensive evaluation of the effectiveness and relevance of each method. The details of the systematic selection, qualitative and quantitative analyses will be discussed below. A full overview of these evaluation steps can be found in Figures 4.9 and 4.10

4.2 Dataset

The research dataset has been gathered by the gynecology department at MMC in Eindhoven and Veldhoven, covering the period from January 18th, 2008, to April 29th, 2021. It consists of a variety of features of 1039 patients, including process-related features such as preoperative and perioperative details. Preoperative features include patient characteristics like age and BMI, as well as sterilization status. Additionally, perioperative features highlight factors occurring during an intervention surgery, such as the type

of anesthesia administered. Furthermore, the dataset includes process features, such as complications on the day the surgery was performed and data collection of the gynecology activities up to 3 years after the surgery.

4.2.1 Preprocessing of the Dataset

The focus of this research is on the process flow before and after an intervention surgery to discover process patterns related to the need for intervention after the surgery. Therefore, the original dataset of MMC has been adapted to perform the research. The MMC dataset consists of different Excel sheets of event logs before and after an intervention, also known as the process features. By merging the variety of event logs within the MMC dataset, one event log was created for the research, which consists of the process features. This resulted in an event log which consist of exactly 1000 patients. This preprocessing step was performed by S. (Suhwan) Lee MSc from the department of Information and Computing Science of Utrecht University.

In Table 4.1, an overview is provided of the different event logs with their descriptions that were merged together. Within this table, various activities are defined, including gynecological care, gynecological orders, and operation activities.

Gynecological care activities are related to the patient's care, such as performing an ultrasound or conducting a telephone consultation. Gynecological order activities involve orders sent by doctors and nurses of the hospital themselves for the patient, such as cervical smear requests for testing purposes and polyclinic laboratory requests for the patient. Lastly, operation activities are associated with the type of surgery performed on the patient, such as sterilization or hysterectomy due to complications.

Table 4.1: The merged event logs and their descriptions

Before or after an intervention	**Description of the event logs**
Before	Collection of complications on the day of surgery.
Before	Description of the appointments 2 years pre-surgery.
Before	Description of gynecological care activities 2 years pre-surgery.
Before	Description of gynecological order activities 2 years pre-surgery.
Before	Description of operation activities 2 years pre-surgery.
After	Collection of complications up to 42 days post-surgery.
After	Description of the appointments 3 years post-surgery.
After	Description of gynecological care activities 3 years pre-surgery.
After	Description of gynecological order activities 3 years pre-surgery.
After	Description of operation activities 3 years pre-surgery.

The merging of the different event logs resulted in a dataset that will be used for cross-validation. The columns and their corresponding data types in this dataset can be viewed in Table 4.2. Furthermore, a statistical overview of this merged data set can be found in Table 4.3.

Table 4.2: Resulting dataset with its column names and its respective data types

Column name	Datatype
CaseID	String and integer
Activity	String
Timestamp	Date (yyyy-mm-dd)
Grouped activity	String
Activity Tag	String
Intervention	Integer (0 or 1)
Data_collection_Appointment_up_to_2_years_-for_Novasure_Appointment_status	String
Data_collection_Appointment_up_to_2_years_-for_Novasure_Appointment_Location	String
Data_collection_gyn_care_activity_up_to_2_-yrs_for_novasure_procedure_care_activity_-code	Integer
Data_collection_gyn_care_activity_up_to_2_-years_for_novasure_procedure_tags	String
Data_collection_gyn_care_activity_up_to_2_-years_for_novasure_procedure_Department	String and integer
Data_Collection_Appointment_up_to_3_-years_after_Novasure_Appointment_status	String
Data_Collection_Appointment_up_to_3_-years_after_Novasure_Appointment_Location	String
Data_Collection_Gyn_Care_Activity_up_to_3_-years_after_Novasure_Proced_Care_Activity_-Code	Integer
Data_Collection_Gyn_Care_Activity_up_to_3_-yrs_after_Novasure_Procedure_Department	String and integer

Table 4.3: Descriptive statistical information of the merged event log

Statistical metric	**Amount**
Events	24 388
Cases	1000
Distinct activities	467
Median case duration	45,2 weeks
Mean case duration	15,6 months
Minimal trace length	2 events
Maximal trace length	149 events
Average trace length	24 events
Median trace length	18 events
Variants	995

4.3 Systematic Selection of Process Patterns

The following paragraphs outline the steps for systematically selecting a total of 24 process patterns from the process pattern discovery of the three PPD methods.

Firstly, A systematic approach for selecting the discovered process patterns for each PDD method will be performed. Subsequently, the qualitative and quantitative analyses will be performed. An overview of the steps for the systematic selection can be found in Figure 4.1.

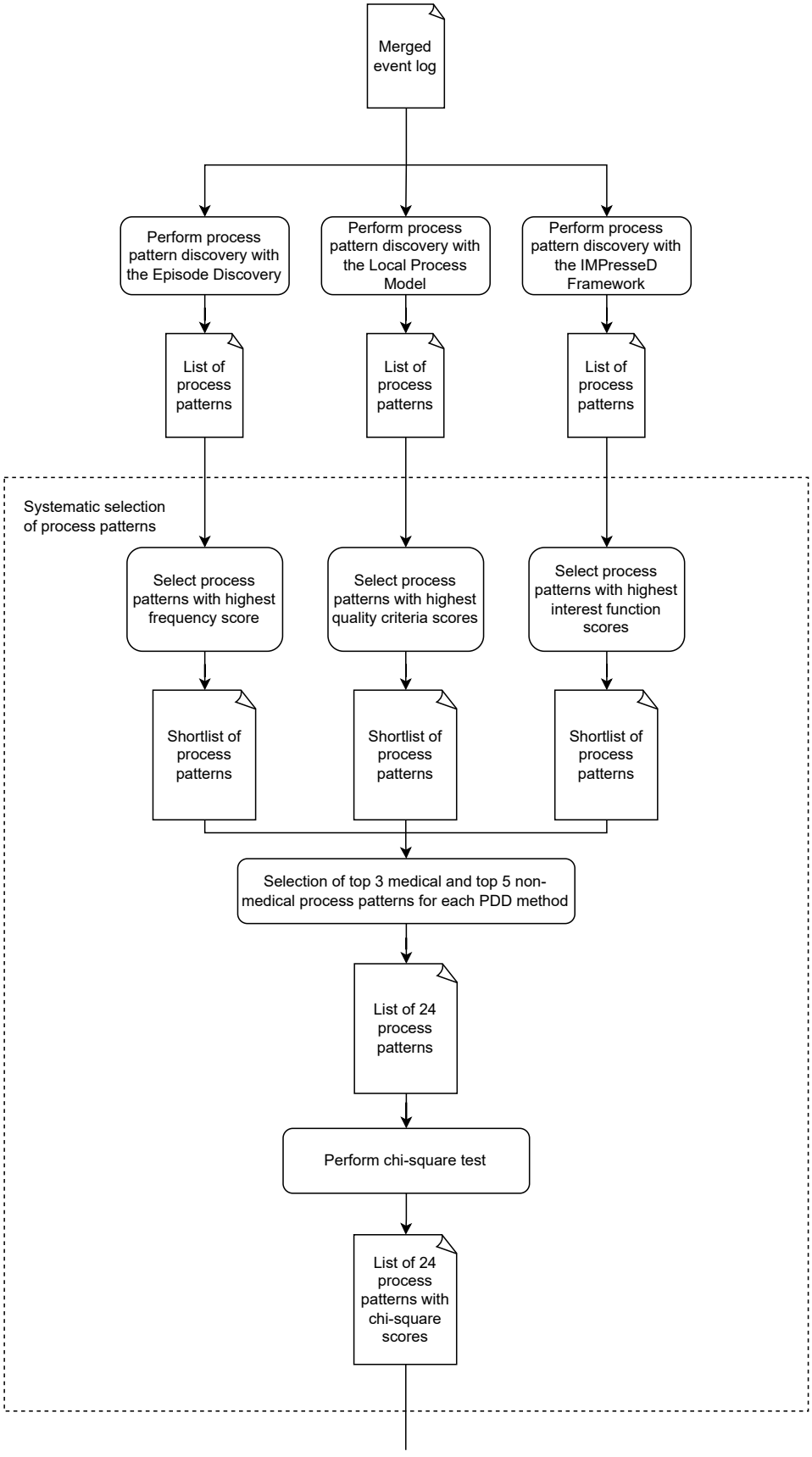

Figure 4.1: Overview of the steps for the systematic selection

The reason for a systematic approach is because due to different type of events within the event log. The type of events can be categorized in *medical* and *non-medical* events. As the names suggest, a medical event is a medical activity such as blood loss, injections or a medical procedure being performed and a non-medical event is an activity that is of non-medical nature such as administration work for the patient or a visit of the patient to the clinic. In Table 4.4 a few examples of medical and non-medical events are listed.

Table 4.4: Example of medical and non-medical events within the event log

Medical events	**Non-medical events**
Before_Echo gynecology via 1st line	Before_First outpatient clinic visit
Before_Ultrasound of the abdominal organs	Before_Control patient gynecology
After_Injection by assistant outpatient clinic	After_Nursing day
Before_Abnormal blood loss	After_Hospitalization

Logically, medical events should have more impact on the outcome of the process, namely if an intervention occurred or not. However, the non-medical events will be taken into account as well in the process pattern discovery due to several reasons. Firstly, the amount of medical events, and hence the amount of process patterns that contain a medical event, are low compared to the amount of non-medical events. Furthermore non-medical process patterns might have an impact on the outcome too. For example a patient who needs a lot of check-ins before they can enter surgery can be an indication of a higher change of intervention. However, this needs to be checked with the medical expert, hence the qualitative analysis.

A systematic approach will be used. First a top three of medical process patterns and a top five of non-medical process patterns will be discovered according to the respective PDD method. Due to the low disparity of medical process patterns only a top three will be used, while for the non-medical process patterns a top five will be discovered.

The definition of top three and top five for each PPD method is determined by how each PDD method discovers its process patterns. For Episode Discovery, the process patterns with the highest frequency will be selected as the top three or top five. For the Local Process Model, the process patterns that perform best overall across all five quality criteria will be chosen for the top three and top five. For the IMPresseD Framework, the top three and top five process pattern will be selected based on how well the perform across all three interest functions. How each PPD method in this research

discovers its process pattern is described in detail in Chapter 3.

In this research a process pattern will be considered *medical* if each event in a process pattern is a medical activity. An example of a medical process pattern is shown in Figure 5.4.

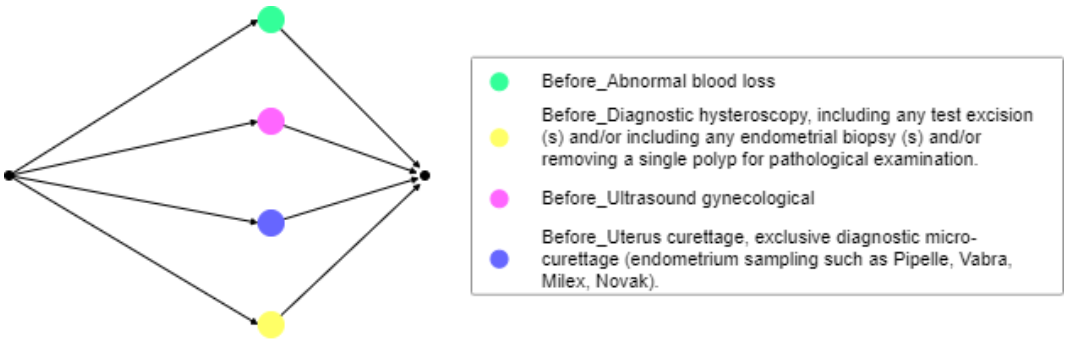

Figure 4.2: A medical process pattern from the IMPresseD framework

A *non-medical* process pattern will consist of a mix non-medical activities and medical activities or only non-medical process patterns. An example of a process pattern that consist of a mix of medical and non-medical events is shown in Figure 4.3.

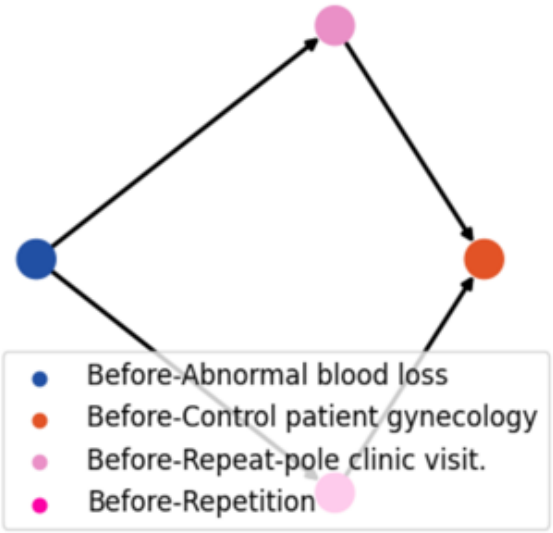

Figure 4.3: A non-medical process pattern from the IMPresseD framework

Eventually, this results in a total of eight discovered process patterns for each method which results in a total of 24 discovered process patterns.

After the selection of the top three and top five process patterns for each

respective method a chi-square test will be performed to calculate the significance of the discovered process patterns to the outcome of the process. This will be done for each discovered process pattern that is in these 24 selected process patterns. The discovered process patterns that are considered significant according to chi-square test will be discussed with a medical expert in the qualitative analysis. The reason for this is that only one hour is allocated for the interview therefore discussing all 24 discovered process patterns will be unrealistic. By using chi-square tests only the significant ones will be discussed to have a fruitful interview.

To perform the chi-square test, several numbers are needed for its calculation. These numbers are extracted from Disco[1], which is a process mining tool designed to analyze and visualize business processes, for the Episode Discovery and the Local Process Model. In Disco the discovered process patterns will be mimicked for the three PPD methods. Using Disco's active filters, like attributes and followers, ensures that specific events are incorporated into a particular flow, with the specification on whether intervention is included or not. For the IMPresseD Framework, the case IDs that contain the discovered process pattern can be extracted from the software itself. However, Disco will be used to extract the cases that contain intervention. This will be compared with the extracted case IDs to perform the chi-square test. These approaches help replicate the discovered process patterns effectively. The numbers that are extracted from Disco represent patients with:

1. Patients with discovered process pattern needing intervention.
2. Patients with discovered process pattern not needing intervention.
3. Patients without discovered process pattern needing intervention.
4. Patients without discovered process pattern not needing intervention.

An example of the extracted numbers from Disco, the IMPresseD Framework, and the chi-square test performed for the discovered process pattern IMP-M1, are shown in Table 4.5. The chi-square p-value is also displayed in the table. Furthermore, the folder structure with the recipes from Disco and the comparison for the IMPresseD Framework can be found in the Appendix D.

[1]https://fluxicon.com/disco/

Table 4.5: Numbers extracted from Disco and the IMPresseD Framework for process pattern IMP-M1 to perform the chi-square test and its p-value

Chi-square statistics	**Intervention**	**No intervention**	**Total**
Pattern detected	2	1	3
Pattern not detected	127	870	997
Total	129	871	**1000**

Chi-square p-value: 0,0012

According to the p-value of the chi-square test in Table 4.5, the discovered process pattern IMP-M1 is significant. To visually highlight other significant discovered process patterns identified by the chi-square test, they will be color-coded in green within their respective tables. The scores for the other chi square tests will be presented in the results chapter.

4.3.1 Episode Discovery

In ProM[2], for both the the Episode Discovery and the Local Process model, the CSV Novasure dataset needs to be converted to a XES format. For the Case column and Event Column select the columns CaseID and Activity respectively from the Novasure dataset. The other settings stay the same, this will result in a XLog file which can be used in the ProM plugins for the Episode Discovery and Local Process Model.

For the Episode Discovery, Leemans Episode Miner - Config dialog by M. Leemans will be used in ProM due to the changes in settings that can be made. This is because due to the low amount of discovered process patterns, hence the frequency threshold will be lowered to 0.2. The other configuration settings stay the same.

As mentioned before, the software Disco will be used to mimic the discovered process patterns to find the case IDs that contain them. For the Episode Discovery the filter follower will be used with a time of less than one day between the selected events that follow each other. This follower filter will be configured to an eventually followed for the selected events. This is due to the settings in ProM for the Episode Discovery. See Figure 4.4 for the described settings in Disco marked in the red boxes.

In addition the several attributes filter will be used to make sure that the events within a discovered process pattern are included in Disco as well,

[2]https://promtools.org/prom-6-13/

see Figure 4.5. Lastly, another attribute filter will be used as well, namely to filter in case IDs that contain intervention or not, see Figure 4.6.

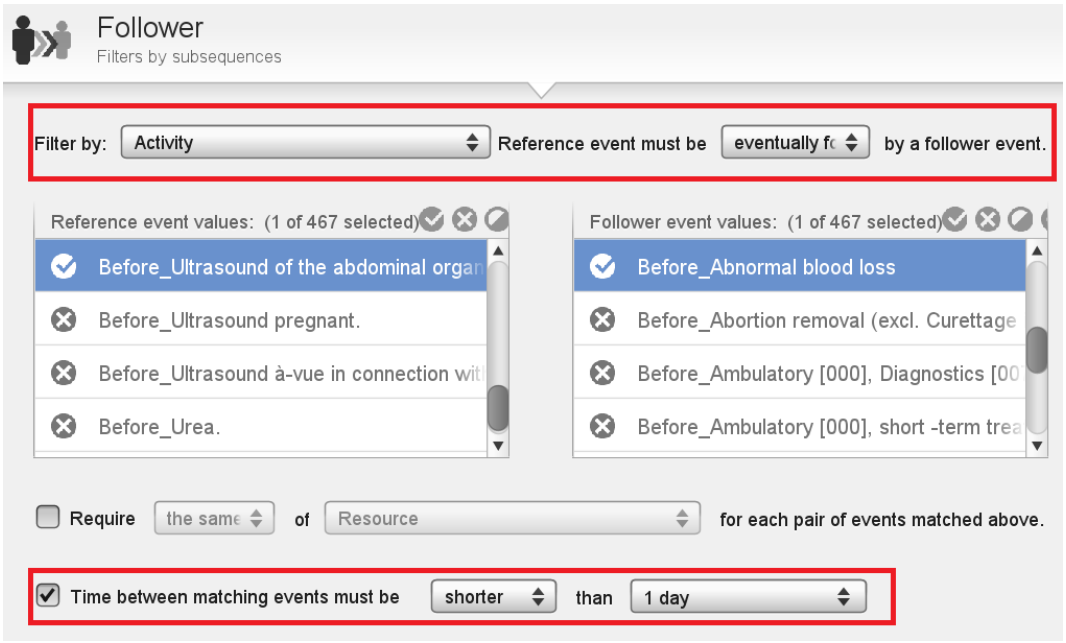

Figure 4.4: In the red boxes the follower filter in Disco are configured as *activity* and *eventually followed*. In addition the time between the follow up events will be *shorter* than *one day*.

The full recipes, which are all the Disco settings used to find the caseIDs of the process patterns, of the Episode Discovery, can be found in Appendix D.

4.3.2 Local Process Model

For the Local Process Model, the Search for Local Process Models in ProM will be used by N. Tax. The reason for using this specific plugins is due to the changes that can be made for discovering local process models. For the grouping of the process patterns Maximal is used because the definition and calculation of Score is unknown so far. The other configuration settings stay the same.

In Disco, the attributes and follower filter were used to ensure that the events in the discovered process patterns were included in the case IDs, directly followed each other, see Figure 4.5, and were filtered based on whether there was an intervention or not, see Figure 4.6.

For the full recipes of the Disco settings for each process pattern of the Local Process Model, see Appendix D.

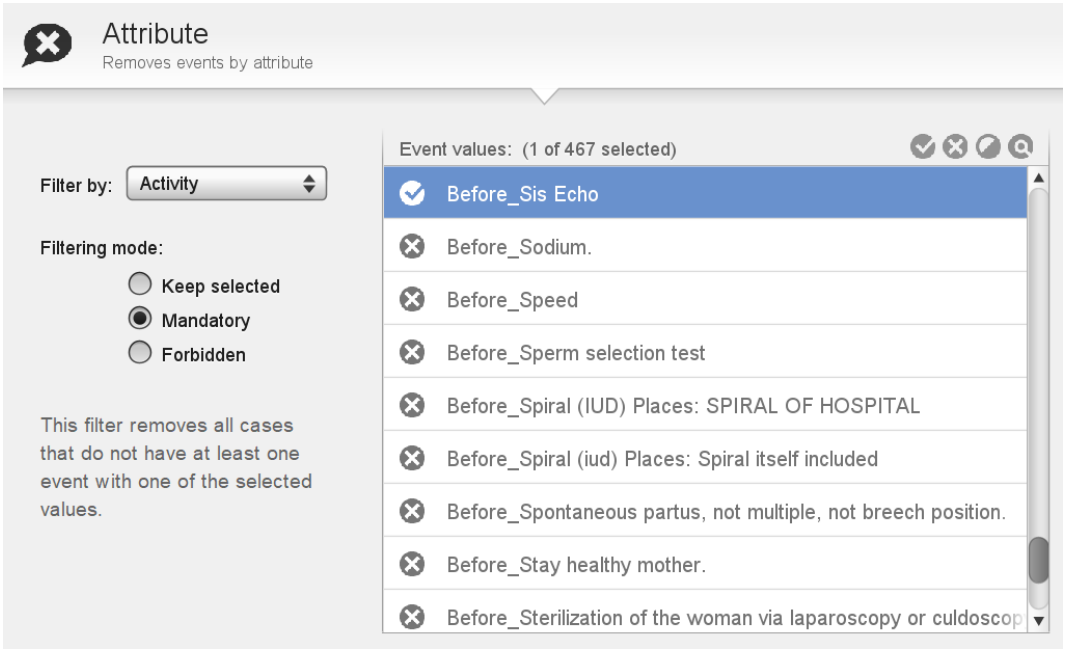

Figure 4.5: The attribute filter in Disco are configured as *Activity* and *mandatory*. In addition, an event is selected to ensure that only the caseIDs will be filtered that contain this event, in this example it is *Before_Sis_Echo*.

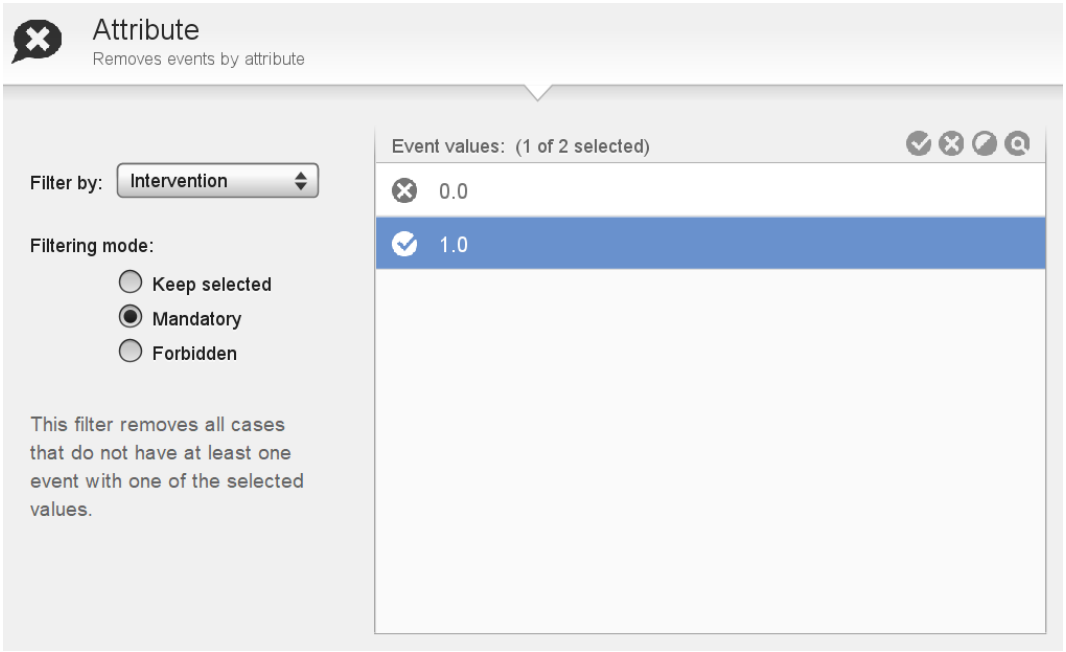

Figure 4.6: The attribute filter in Disco are configured as *Intervention* and *mandatory*. In addition, 1.0 is selected to filter only the caseIDs that contain an intervention.

4.3.3 IMPresseD Framework

The IMPresseD Framework has its own software on GitHub[3], by Mozhgan Vazifehdoostirani from Eindhoven University of Technology, to discover process patterns. The configurations settings for the IMPresseD Framework for the Novasure dataset are the following:

- Case ID column: CaseID
- Activity column: Activity
- Outcome column: Intervention
- Outcome type: binary
- Timestamp column: Timestamp
- Delta time (in seconds): 1
- Categorical attributes: none
- Numerical attributes: none
- Visualization row number: 2
- Visualization column number: 1
- Case Distance interest function: Checked & Min
- Frequency interest function: Checked & Max
- Correlation interest function: Checked & Max

For the Case Distance interest function, a minimum score is preferred because, ideally, research aims to limit the effect of confounding variables. For the Frequency interest function, a maximum score is preferred because the more often a process pattern occurs, the more crucial it is likely to be within the process. Similarly, for the Correlation interest function, a maximum score is also preferred, as it indicates the extent to which the discovered process pattern impacts the outcome column, which in this case is Intervention. With the above configuration settings, the Interactive Pattern Discovery button can be clicked. This will present several process patterns and their outcomes for the checked interest functions.

As mentioned before, to calculate the chi-square for the discovered process patterns of the IMPresseD Framework, the case IDs will be extracted from the IMPresseD Framework software itself. Furthermore, the case IDs

[3]https://github.com/MozhganVD/InteractivePatternDetection

that contain intervention will be extracted from the dataset via Disco. By comparing the case IDs that contain intervention from Disco with the case IDs that contain a specific discovered process pattern from the IMPresseD Framework software, the numbers needed for the chi-square test can be extracted.

4.4 Qualitative Analysis

The following paragraphs describe the procedure for performing the qualitative analyses, namely the selection process for the three process patterns for each PDD method, which results in a total of nine process patterns, and the setup of the interview.

For the qualitative analysis, an interview will be arranged with a medical expert from the MMC in Eindhoven. The aim of this interview is to facilitate an in-depth discussion regarding the findings of the three PPD methods with a medical expert. This is to validate the quality of the discovered process patterns with the medical expert's knowledge. The quality will be checked by asking the medical expert how crucial a process pattern is with regards to intervention. This is done by checking the *alignment* of the process patterns with a medical expert. In this research setting, alignment means how much the order and combination of events within a process pattern is in line with the expectation of a medical expert. Alignment is used to compare the three PPD methods with each other. The PPD method that discovers the process patterns which are most aligned with the medical experts expectation, will be considered as qualitatively better in discovering process patterns in a medical setting. An overview of the steps for the qualitative analysis can be found in Figure 4.7.

As mentioned earlier, three process patterns of each PDD method will be selected from the systematic selection. The aim is to provide a balanced representation by including both medical and non-medical process patterns. Consequently, the interview lineup will feature a mix of significant and non-significant process patterns, as well as both medical and non-medical ones. This approach ensures that from the 24 discovered patterns, the most compelling nine are selected for the interview.

The evaluation objective in this analysis is to assess the alignment of the discovered process patterns with a medical expert expectation and knowledge. To assess the alignment, a medical expert's expectations on how each

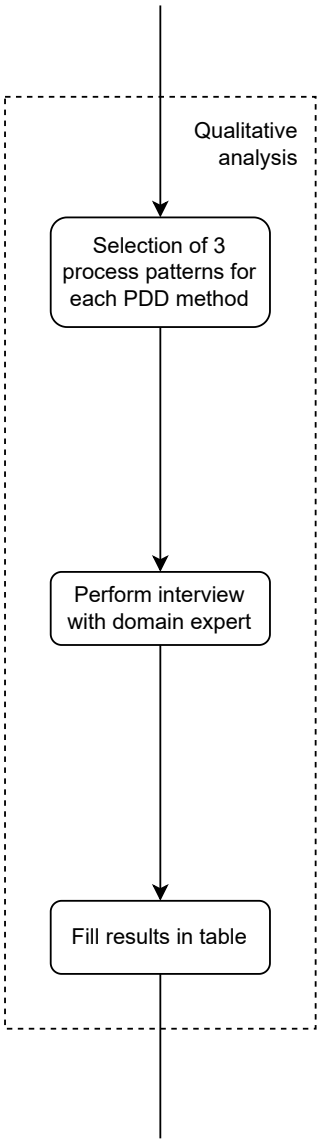

Figure 4.7: Overview of the steps for the qualitative analysis

process pattern relates to intervention are compared before and after presenting the chi-square test results.

By comparing the results, a better understanding of the quality of the three PPD methods will be acquired, and the reasons behind these differences will be explored. The answers for each process pattern will be documented in a form. Additionally, the interview session will be recorded. Due to time constraints, only nine process patterns will be discussed. The type of question and the questions itself are presented below:

1. **Closed question**: Is this a process pattern you would expect to be important for the outcome according to your knowledge?

2. **Linear scale question (1-10)**: For the question above, how confident are you?
3. **Open question**: For the question above, why?
4. **Closed question**: Did you expect this (after showing chi-square test result)?
5. **Closed question**: For the question above, why?

Lastly, at the end of the interview, both a closed and an open question will be asked about the representation of the models produced by the three PPD methods. Specifically:

1. **Closed question**: Below you see three traces of three patients with their models. Which representation do you prefer?
2. **Open question**: For the question above, why?

Outlined below is the structure of the interview. Initially, the significance of process patterns, as determined by chi-square test, is withheld to prevent potential bias in the responses of medical expert. This decision aims to elicit their unbiased opinions and genuine expectations regarding the process patterns. The significance will only be revealed after their answer about the potential importance of a given process pattern in relation to intervention. This enables a comparative assessment of the nine process patterns, ensuring consistency by using standardized questions for each process pattern. Ultimately, the objective is to gain insights into the quality of process pattern selection algorithms employed by different PPD methods. Each process pattern will follow the same structure during the interview:

1. Introduction to the three PPD methods utilized in the research, focusing on their process pattern discovery. This will mostly be done via email to make sure the medical expert know what to expect and prepare for the interview. Hence, this step will only be performed once.
2. Show the process pattern and ask the medical expert if they think this process pattern is crucial for the outcome, with regards to intervention or not.
3. Show if it was significant according to the chi-square test.
4. Discuss whether the medical expert expected this or not, why?
5. Summarizing the key points of this process pattern.

These results will then be used to answer the research question *To what*

extent are the discovered process patterns aligned with the medical experts expectations? as described in section 2.

4.5 Quantitative Analysis

The following paragraphs describes in detail the pipeline of the quantitative analysis, including the measurements and models that will be used.

For the quantitative analysis, the predictive capabilities of the discovered process patterns will be assessed. Specifically, this assessment involves evaluating their ability to predict potential re-interventions after an intervention surgery using the evaluation objective of accuracy, F1 and AUC-ROC value. An overview of the steps for the quantitative analysis can be found in Figure 4.8.

To accomplish this, an additional feature indicating the presence of a process pattern will be added to the event log on a case level. This feature will be encoded as 0 if the case does not contain the discovered process pattern and 1 otherwise. Moreover, the prediction task will focus on determining whether an intervention will occur based on the presence of the identified process pattern in the case. The intervention status will also be represented as 0 or 1 in the datasets.

To conduct this prediction, cross-validation will be employed, a technique used to train and validate predictive models while quantifying their performance. This approach will yield accuracy, F1, and AUC-ROC values for each PPD method and its associated process patterns. A higher accuracy score indicates better predictive ability regarding the necessity of intervention based on the identified process patterns. Similarly, a higher F1 score reflects a balanced precision and recall, demonstrating the model's effectiveness in identifying true positives and minimizing false positives and negatives, which is especially important in the medical field. The AUC-ROC value, representing the area under the receiver operating characteristic curve, measures the model's ability to distinguish between classes, with a higher AUC-ROC value indicating a superior performance in differentiating between instances where intervention is needed and those where it is not. Furthermore, the AUC-ROC curve is also less affected by imbalanced datasets which is the case in this research and comes often a lot in medical settings. Thus, the process patterns with relatively higher accuracy, F1, and AUC-ROC values can be considered critical for needing an interven-

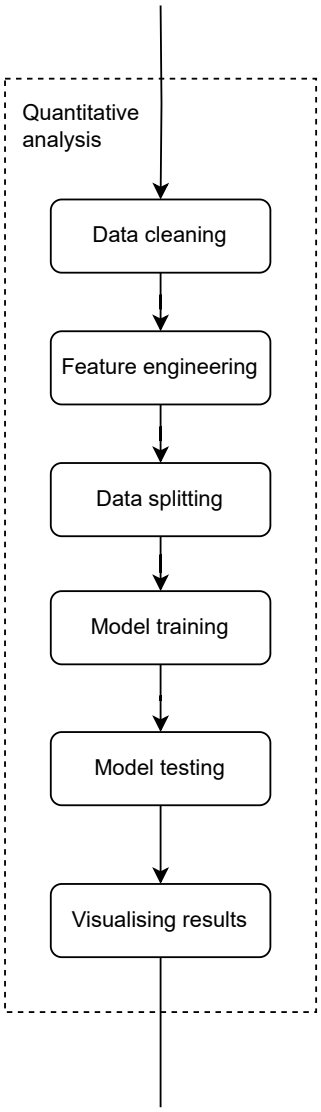

Figure 4.8: Overview of the steps for the quantitative analysis

tion. Consequently, the respective PPD method can recognize and extract these critical process patterns from an event log.

All 24 process patterns will undergo cross-validation. Each process pattern will be evaluated individually, resulting in three separate evaluations for each PPD method.

The pipeline for cross-validation is outlined below, with each step numbered for clarity:

1. **Data Cleaning**: The dataset will be cleaned by organizing it into separate columns in Excel, with any null or incomplete values replaced with "none." The dataset will include columns for caseIDs, the top 10 most frequent activities, intervention, and process patterns.

The top 10 most frequent activities and their frequencies will be extracted from Disco and can be viewed in Table 4.6. The reason for focusing on the top 10 activities is to enhance the model's learning capability. Given the limited number of features, which are caseIDs, process patterns, and intervention, incorporating the top 10 frequent activities adds valuable information for the model, thereby enhancing its predictive performance.

All columns, except for caseIDs, will be encoded as 0 or 1. The caseIDs will consist of a combination of numbers and letters. Furthermore, each of the top 10 most frequent activities will have its own column. If a caseID includes an activity, it will be encoded as 1. This results in a total of 13 columns and 1,001 rows.

2. **Feature Engineering**: A feature indicating the presence of a process pattern will be engineered and added to the dataset. This feature will be encoded as 0 or 1, where 1 denotes the presence of the process pattern and 0 denotes its absence. The presence of each process pattern in a caseID will be extracted from Disco and incorporated into the dataset, aligning with the corresponding caseIDs.

3. **Dataset Splitting**: The dataset will be divided into 10 folds for cross-validation. In each iteration, 9 folds will be used for training the model, while the remaining 1 fold will be used for validation. As a result, each fold will contain approximately 100 caseIDs for training or testing.

4. **Model Training**: From the data splitting, 9 folds will be used to train the predictive models. Specifically, logistic regression and decision tree classification algorithms will be employed. Logistic regression is ideal for binary classification tasks, as it models the probability of a binary outcome based on input features. Which is ideal for these datasets, because the input and output features are all binary encoded.

   In contrast, decision trees are useful for binary data because they can capture complex interactions between features through a series of hierarchical decision rules. Decision trees can handle binary features well and can reveal how combinations of features influence the outcome by splitting the data into branches based on feature values. This is specifically useful in this research because a combination of activities can have an impact on the outcome, namely intervention or not. Hence it is important to capture this complex interaction.

   Additionally, to assess potential improvements, the models will first

be trained without the process pattern column to use as a baseline. Subsequently, a new model will be trained with the process pattern column included. The performance metrics for the models with and without the process pattern column will be compared to evaluate how effectively the PPD method captures important process patterns related to intervention.

Lastly, a grid search will be used to address the class imbalance, as there are only 129 cases with intervention out of 1,000.

5. **Model Testing**: The trained model will be evaluated on the remaining fold from the data splitting to calculate accuracy, F1 score, and AUC-ROC value. In this research, the intervention column is the prediction target, while the top 10 frequent activities and process patterns are the input features.
6. **Visualising results**: Accuracy, F1, and AUC-ROC values will be visualized using graphs and tables to track potential improvements. The average metrics for the selected process patterns for each PPD method will be used to compare the results with the baseline, namely the trained model without the process pattern column.
7. **Analysing results**: The obtained results will be presented in the Chapter 5. Furthermore, it will be analysed and discussed in Chapter 7 and 6.

Overall this quantitative analysis, based on the resulting accuracy, F1, and AUC-ROC values from the validation set of the cross-validation, should provide valuable insights into the effectiveness of the patterns in predicting post-surgical interventions. This will be used to answer the research question *Performance-wise, what are the differences in the discovered process patterns across various process pattern discovery methods, in terms of accuracy?* as described in Chapter 2. By answering this research question there will a better understanding of the differences between the various PPD methods.

An overview of the complete evaluation setup is provided in Figures 4.9 and 4.10

In addition to cross-validation, feature importance will also be performed, using the decision tree model. This is performed to determine which process patterns, and consequently which PPD methods, have the most influence on the outcomes of a potential re-intervention. The decision tree model was chosen over the logistic regression model because it is an open-box model,

Table 4.6: Top 10 most frequent activities with its frequency

Activity	Frequency
After_Consult by phone	1539
Before_Ultrasound of the abdominal organs.	1524
Before_First outpatient clinic visit.	1177
After_Repeat-pole clinic visit.	1171
Before_Repeat-pole clinic visit.	1073
Before_Consult by phone	965
Before_Abnormal blood loss	913
After_Teleconsult (excluding screen to screen image contact, see 190019).	892
Before_First outpatient visit	879
Before_Repetition	870

meaning its decisions can be traced and understood, unlike the black-box nature of logistic regression. This transparency is crucial in a medical setting where the decisions made by a machine learning model need to be explainable and supported by clear reasoning. Therefore, the decision tree model will be used for the feature importance analysis.

This analysis will be performed using the top 10 events mentioned in Table 4.6, along with the caseID and intervention columns. This first analysis will be used as a baseline. Next, the 24 process patterns will be added to the top 10 events, creating a total of 36 columns in the dataset, including the caseID and intervention columns. Additionally, another analysis will be performed using only the 24 process patterns. This results in a total of three analyses and three datasets: the top 10 events from Disco, the top 10 events plus the 24 process patterns, and just the 24 process patterns. These datasets and Jupyter Notebook code can be found in the Appendix D.

4.6 Threats to validity

While this evaluation plan is thoughtfully constructed, limitations still exist within this plan. These limitations will be described in the following paragraphs.

The systematic selection of process patterns has limitations. Specifically, it is done manually based on overall best performance, depending on the discovery technique used for each PPD method, and whether the process pattern is *medical* or *non-medical*. This approach could potentially lead to biases because the selection is not performed by a domain expert, which could result in choosing process patterns that are not interesting due to a lack of medical knowledge. Furthermore, the overall best performing process patterns, according to the respective PPD method, mostly consisted of only two events. Therefore, a minimum of three events per process pattern was also included in the systematic selection. However, this resulted in lower scores for the frequency, quality criteria, and interest functions for the process patterns that included at least three events. Hence, process patterns with two events are also included.

Due to potential time constraints during the interview, in-depth discussions of each PPD method are limited to just three patterns. Therefore, patterns that are not discussed during the interview may be overlooked, but could potentially be crucial. In addition, time constraint may pose a challenge during the interview itself. Although thorough preparation is assumed, time limitations may prevent a complete discussion of all the nine patterns. Limitations could impact the validity of the qualitative analysis.

Limitations also exist in the quantitative analysis. During this analysis, more than three patterns will be utilized in cross-validation. However, because only three patterns for each PPD method are discussed during the interview, the validity of the quantitative analysis is also diminished. Patterns not discussed during the interview will still be used in the quantitative analysis, thereby lowering its validity. Ideally, all the patterns used during the quantitative analysis should also be discussed during the qualitative analysis. However, due to time constraints during the interview, this is not possible.

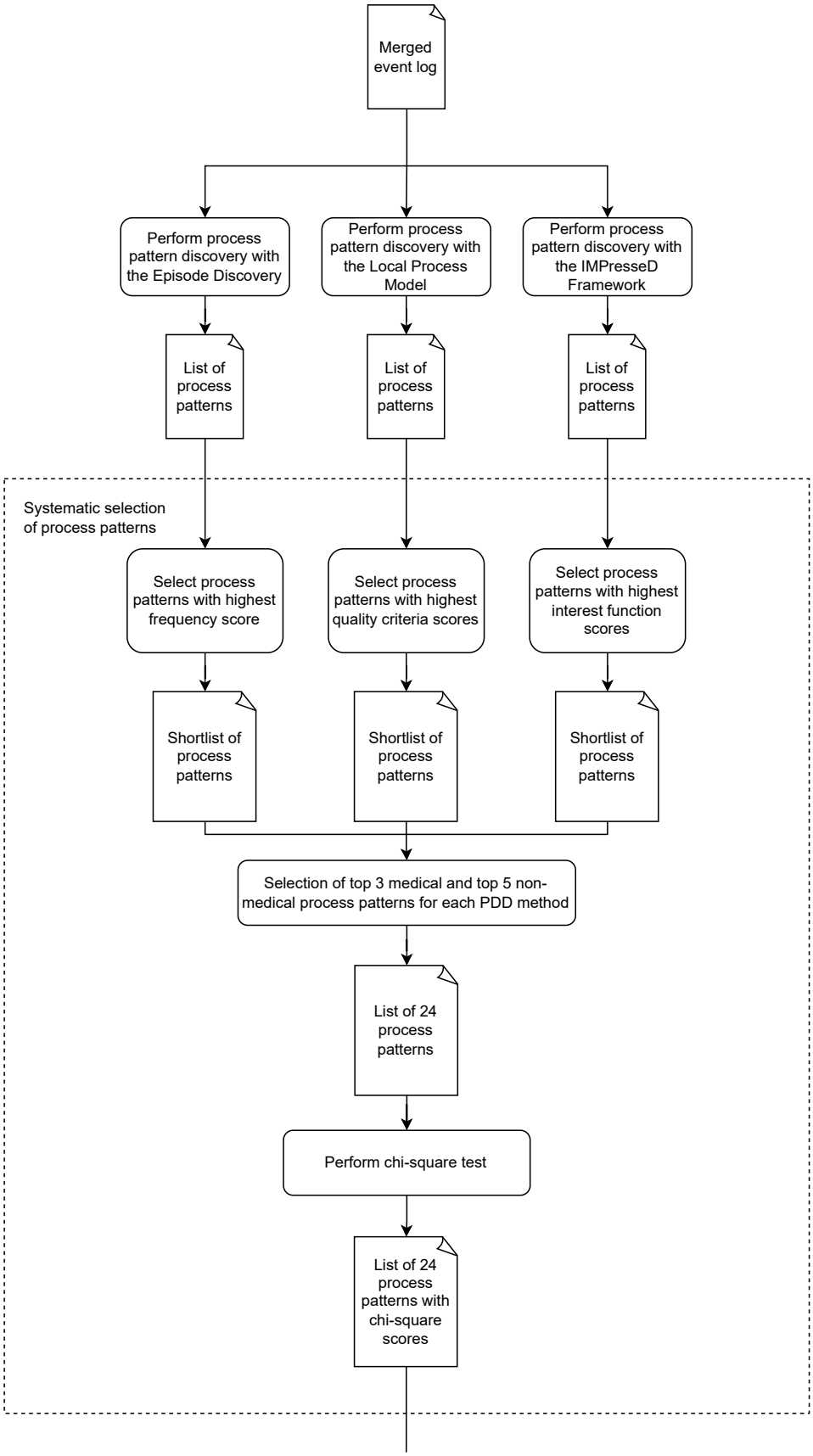

Figure 4.9: First part of the steps of the evaluation setup

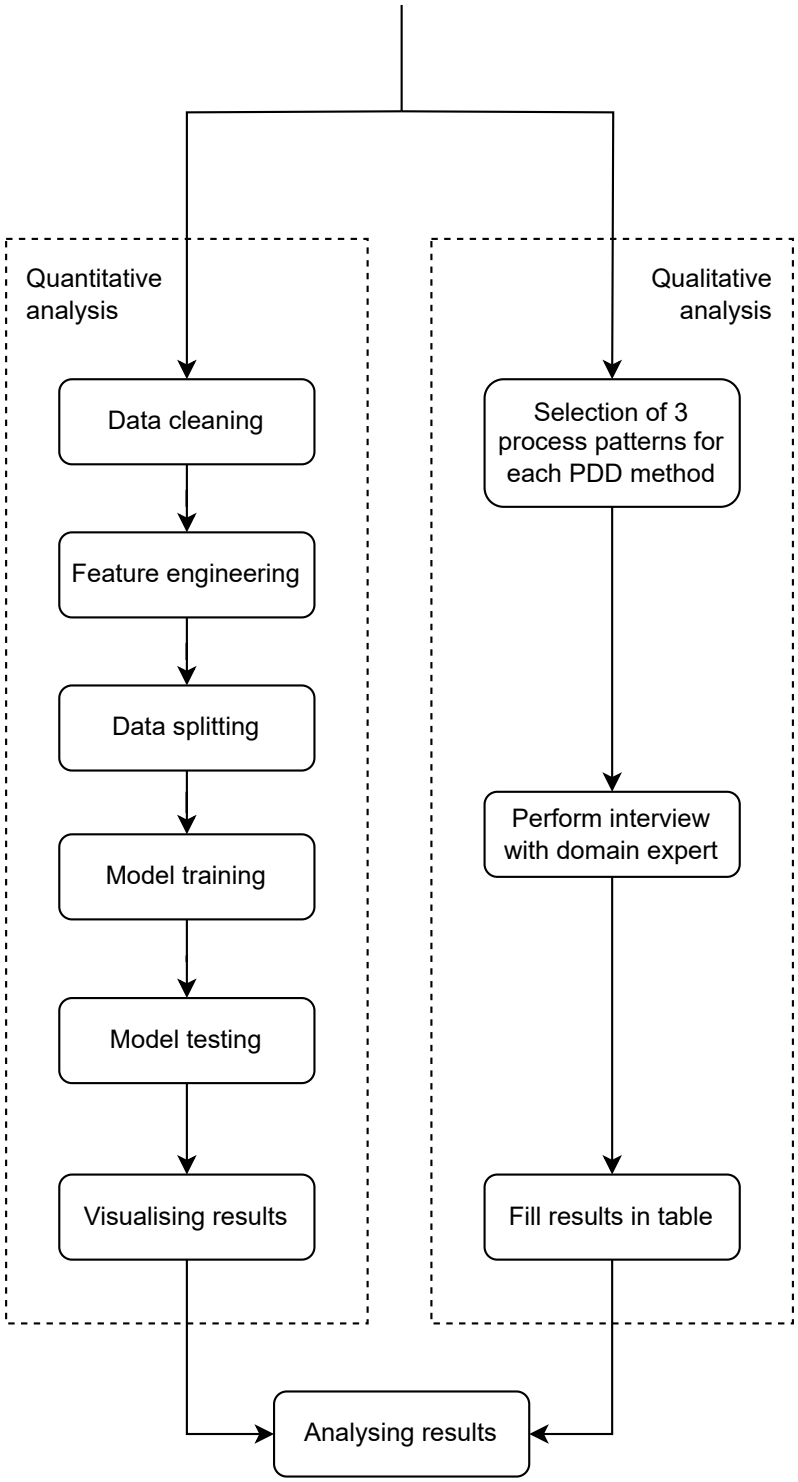

Figure 4.10: Second part of the steps of the evaluation setup

5. Results

In this chapter the results of the systematic selection, qualitative analysis, and quantitative analysis will be presented. The outcomes of the systematic selection will then be used in the qualitative analysis to make a selection for the interview with the medical expert.

5.1 Systematic Selection and their Significance

In the following sections the results for each PPD method will be presented and discussed. These results consists of the selection of the top three medical and top five non-medical process patterns, and the chi-square scores.

5.1.1 Episode Discovery

After the systematic selection for the Episode Discovery, the top three medical process patterns can be found in the following Table 5.1. The visual representation of these patterns can be found in the Appendix A.1. For naming and readability purposes, in this research, the first letters before the hyphen are the abbreviation of the PPD method, in this example Episode Discovery (ED), and the letters after the hyphen stand for medical (M) or non-medical (NM). This holds for the other two PPD methods as well.

Table 5.1: Top three medical process patterns for Episode Discovery and their frequency.

Top three process patterns	Frequency
ED-M1	0,170
ED-M2	0,179
ED-M3	0,229

Furthermore, the top five discovered non-medical process patterns for the Episode Discovery can be found in Table 5.2. The visual representation of these patterns can be found in the Appendix A.1.

An initial look at the top three medical and top five non-medical process patterns reveals that the frequency of the top five non-medical patterns is higher than that of the top three medical ones. This is understandable because purely medical process patterns, consisting only of medical events, are relatively rare in the list of discovered process patterns. As a result, the

Table 5.2: Top five non-medical process patterns for Episode Discovery and their frequency.

Top five process patterns	Frequency
ED-NM1	0,331
ED-NM2	0,309
ED-NM3	0,296
ED-NM4	0,275
ED-NM5	0,225

frequency of medical process patterns is quite low, indicating that there are not many of them in the event log.

In addition, in Table 5.3 the numbers for the calculation of the chi-square test, including the chi-square value, are presented for the selected process patterns from the Episode Discovery. Marked in green are the process patterns that are considered significant according to the chi-square test.

Table 5.3: Table with the numbers for the calculation of the chi-square, including the chi-square value, for the selected process patterns for the Episode Discovery.

	Detected		*Undetected*		
Selected process patterns	**Re-intervention**	**No re-intervention**	**Re-intervention**	**No re-intervention**	**Chi-square**
ED-M1	19	151	110	720	0,461810762
ED-M2	41	273	88	598	0,920015313
ED-M3	29	281	100	590	0,024977137
ED-NM1	60	404	69	467	0,978267808
ED-NM2	62	407	67	464	0,776888148
ED-NM3	39	342	90	529	0,048659075
ED-NM4	39	239	90	632	0,508752317
ED-NM5	31	230	98	641	0,56642351

Table 5.3 shows that two process patterns from the Episode Discovery are significant according to the chi-square test: ED-M3, a medical process pattern, and ED-NM3, a non-medical process pattern. These process patterns are visible in Figure 5.1 and Figure 5.7. However, a deeper analysis of the expected and actual frequencies for these chi-square tests reveals that the expected frequencies is higher than the actual frequencies for both process patterns. This discrepancy is why they are considered significant by the chi-square test. This means that the reason why they are significant are because the actual frequencies, visible in Table 5.3, are lower compared to the expected frequencies. Ideally, the actual frequency should be higher than the expected frequencies, hence why they are significant but that is not the case here.

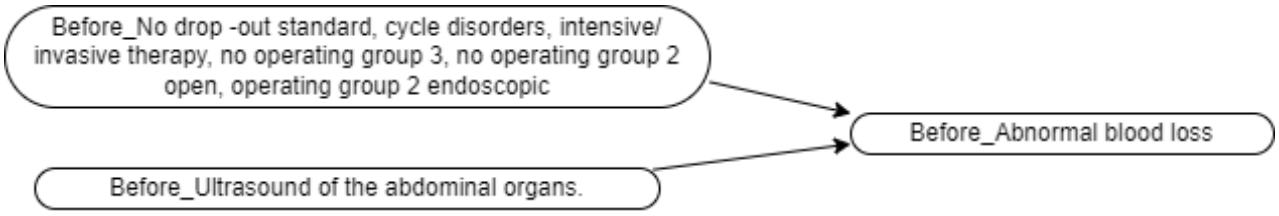

Figure 5.1: Process pattern ED-M3, which is a medical process pattern from the Episode Discovery

Figure 5.2: Process pattern ED-NM3, which is a non-medical process pattern from the Episode Discovery

5.1.2 Local Process Model

The top three discovered medical process patterns for the Local Process Model can be found in Table 5.4. The top five discovered non-medical process patterns can be found in Table 5.5. The visual representation of these patterns can be found in the Appendix A.2.

Table 5.4: Top three medical process patterns for Local Process Model and their quality criteria.

Top three process patterns	Support	Confidence	Language fit	Determinism	Coverage
LPM-M1	13	0,70	1	0,66	0,0015
LPM-M2	5	0,83	1	1	0,000492
LPM-M3	7	0,57	0,75	0,82	0,003

Table 5.5: Top five non-medical process patterns for Local Process Model and their quality criteria.

Top five process patterns	Support	Confidence	Language fit	Determinism	Coverage
LPM-NM1	626	0,585	1	1	0,0878
LPM-NM2	55	0,196	1	1	0,034
LPM-NM3	78	0,560	1	1	0,040
LPM-NM4	6	0,308	1	1	0,001
LPM-NM5	5	0,294	1	1	0,001

Similar to the Episode Discovery, the support of the Local Process Model are higher for the non-medical process pattern than for the medical process

patterns. Furthermore, the coverage is also relatively higher for the non-medical than for the medical process pattern. Interestingly, the confidence for the three medical process patterns is higher compared to the top five non-medical process patterns.

In Table 5.6 the numbers for the calculation of the chi-square test, including the chi-square value, are presented for the selected process patterns from the Local Process model. Marked in green are the process patterns that are considered significant according to the chi-square test.

Table 5.6: Table with the numbers for the calculation of the chi-square, including the chi-square value, for the selected process patterns for the Local Process Model.

	Detected		*Undetected*		
Selected process patterns	**Re-intervention**	**No re-intervention**	**Re-intervention**	**No re-intervention**	**Chi-square**
LPM-M1	2	10	127	861	0,695338646
LPM-M2	0	5	129	866	0,388303043
LPM-M3	2	5	127	866	0,214493
LPM-NM1	80	504	49	367	0,372022271
LPM-NM2	11	35	118	836	0,022522703
LPM-NM3	19	83	110	788	0,068600853
LPM-NM4	1	1	128	870	0,117156267
LPM-NM5	1	7	128	864	0,972966507

As highlighted in Table 5.6, the process pattern LPM-NM2, shown in Figure 5.3, is significantly correlated with re-intervention. A closer look at the actual and expected frequencies reveals that the actual frequencies for the detected process pattern and re-intervention are higher than expected. Which is, as mentioned before, ideally the preference.

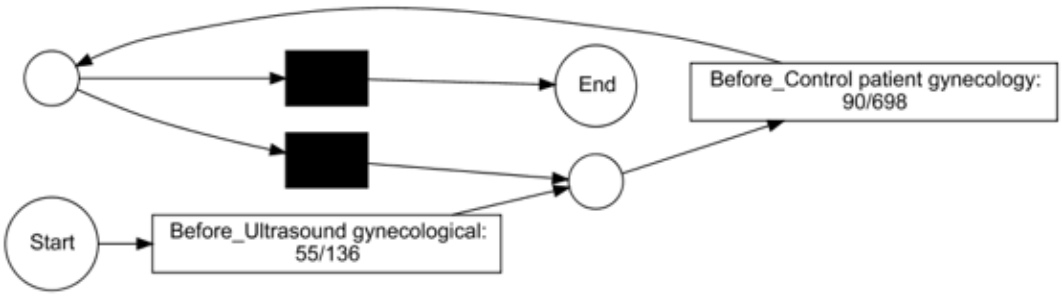

Figure 5.3: Process pattern LPM-NM2, which is a non-medical process pattern from the Local Process Model

5.1.3 IMPresseD Framework

in Table 5.7 the top three discovered medical process patterns for the IMPresseD Framework are presented. The top five discovered non-medical

process patterns are presented in Table 5.8. The visual representation of these patterns can be found in the Appendix A.3.

Table 5.7: Top three medical process patterns for the IMPresseD Framework.

Top three process patterns	Frequency interest function	Outcome interest function	Correlation interest function
IMP-M1	0,003	0,002	0,879
IMP-M2	0,165	0,001	0,798
IMP-M3	0,002	0,001	0,744

Table 5.8: Top five non-medical process patterns for the IMPresseD Framework.

Top five process patterns	Frequency interest function	Outcome interest function	Correlation interest function
IMP-NM1	0,179	0,270	0,820
IMP-NM2	0,139	0,024	0,789
IMP-NM3	0,037	0,035	0,805
IMP-NM4	0,025	0,0034	0,724
IMP-NM5	0,031	0,293	0,848

What stands out when looking at Tables 5.7 and 5.8 is, that the outcome interest functions for the top three medical process patterns are relatively low compared to the top five non-medical process patterns. This indicates that the top three medical process patterns are less correlated with the outcome, whether it be re-intervention or no re-intervention.

In Table 5.9 the numbers for the calculation of the chi-square test, including the chi-square value, are presented for the selected process patterns from the IMPresseD Framework. Marked in green are the process patterns that are considered significant according to the chi-square test.

Table 5.9: Table with the numbers for the calculation of the chi-square, including the chi-square value, for the selected process patterns for the IMPResseD Framework.

Selected process patterns	*Detected*		*Undetected*		Chi-square
	Re-intervention	No re-intervention	Re-intervention	No re-intervention	
IMP-M1	2	1	127	870	0,005395565
IMP-M2	19	146	110	725	0,561401888
IMP-M3	1	1	128	871	0,117156267
IMP-NM1	55	124	74	747	4,07635E-15
IMP-NM2	15	124	114	747	0,424123844
IMP-NM3	7	30	122	841	0,265700599
IMP-NM4	0	25	129	846	0,05132671
IMP-NM5	21	10	108	861	2,16318E-20

As highlighted in Table 5.9, the process patterns IMP-M1, shown in Figure 5.4, IMP-NM1, shown in Figure 5.5, and IMP-NM5, shown in Figure 5.6, are significantly correlated with re-intervention. A closer look at the actual and expected frequencies reveals that the actual frequencies for the detected process patterns and re-intervention are higher than expected. This aligns with the ideal preference mentioned earlier.

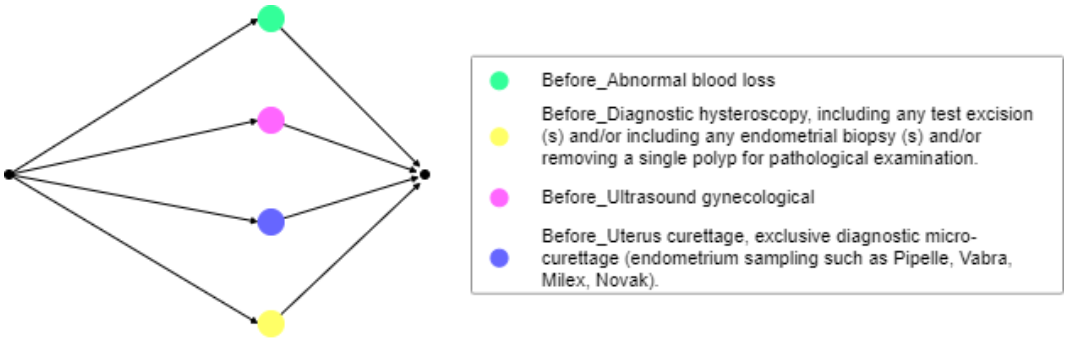

Figure 5.4: Process pattern IMP-M1, which is a medical process pattern from the IMPresseD Framework

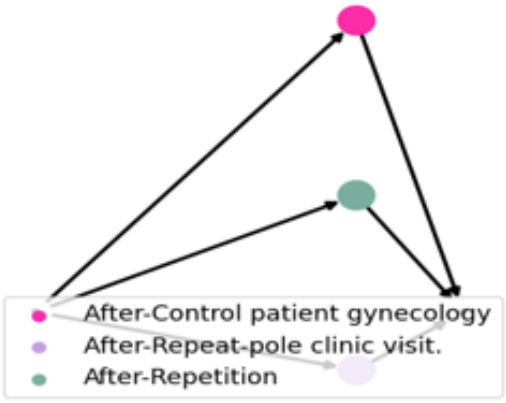

Figure 5.5: Process pattern IMP-NM1, which is a non-medical process pattern from the IMPresseD Framework

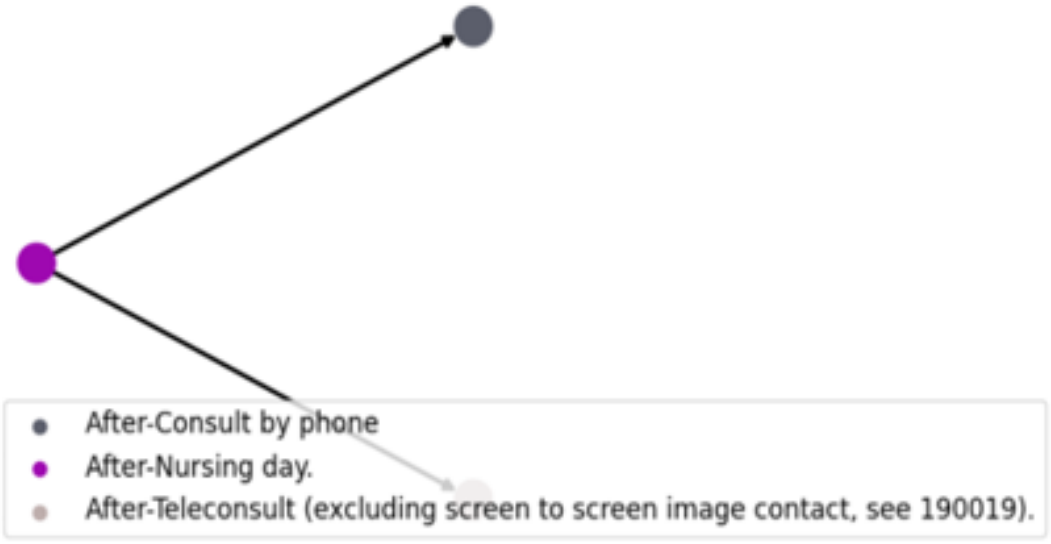

Figure 5.6: Process pattern IMP-NM5, which is a non-medical process pattern from the IMPresseD Framework

5.1.4 Discussion

Based on the results of the systematic selection and chi-square tests, the IMPresseD Framework exhibits the most significant process patterns, specifically three in number. Episode Discovery follows with two significant process patterns. However, these patterns are significant because the actual frequencies are lower than the expected frequencies, which is not ideal; ideally, the actual frequency should exceed the expected frequency. The Local Process Model shows one significant process pattern. Compared to Episode Discovery, the actual and expected frequencies reveal that the actual frequencies for the detected process pattern and re-intervention are higher than expected, which is, as mentioned earlier, the preferable outcome.

Upon inspecting the process patterns, it was found that the patterns IMP-M2 and ED-M1 are identical, having been discovered by both PPD methods. All three PPD methods share similar events, such as "Before_Control patient gynecology," "Before_First outpatient clinic visit," and "Before_Consult by phone." Additionally, Episode Discovery identified process patterns with the same events but in a different order, such as ED-NM1 and ED-NM3. Furthermore, ED-NM4 and IMP-NM3 are almost identical, except that IMP-NM3 includes an additional event: "Before_Abnormal blood loss," which precedes the other three events.

Notably, many process patterns discovered in the IMPresseD Framework occurred after the intervention. In fact, two of these process patterns, IMP-M1 and IMP-NM5, are significant. Additionally, the Local Process Model includes process patterns that consist of two events and loops, with one of these loops, LPM-NM2, being significant. This is also the only significant process pattern for the Local Process Model.

5.2 Qualitative Analysis

5.2.1 Episode Discovery

For the interview a selection is made to discuss during the interview. For the Episode Discovery the selection can be found in Table 5.10. The significant process patterns are highlighted in green based by the chi-square test.

Table 5.10: Selected process patterns of the Episode Discovery, which will be discussed during the interview.

Top three process patterns	Frequency
ED-M3	0,229
ED-NM1	0,331
ED-NM3	0,296

Furthermore, in Table 5.11 the results of the interview questions with the medical expert can be found.

An initial examination of these results reveals that the significant process patterns from the Episode Discovery, specifically ED-M3 and ED-NM3, are not expected to be important for the outcome of re-intervention or no re-intervention. For ED-M3, one of the events was unclear for the medical expert due to multiple factors being crammed into it. Regarding ED-NM3, the medical expert explained that while the activities in the process pattern ED-NM3 are reasons why a patient might undergo an intervention, they are not necessarily indicative of an re-intervention. Thus, the medical expert does not consider this pattern important for the outcome. However, when showing the chi-square score, the medical expert noted that literature shows that 10% of patients who undergo intervention need a re-intervention due to ongoing complaints.

The findings for ED-NM1, shown in Figure 5.7, are also quite interesting. According to the medical expert, the activities are not in the correct order. From the patient's perspective, the order is first the complaint which is abnormal blood loss, then a first visit to the clinic, followed by an ultrasound of the abdominal organs. However, the doctor records these events in a different order in the system.

Figure 5.7: Process pattern ED-NM3, which is a non-medical process pattern from the Episode Discovery

Table 5.11: Qualitative analysis for the Episode Discovery.

Questions	ED-M3	ED-NM1	ED-NM3
Is this a process pattern you would expect to be important for the outcome according to your knowledge?	No	No	No
For the question above, how confident are you? (Scale from 1-10)	4	9	9
For the question above, why?	The first activity is unclear because a variety of factors have been crammed into it.	The sequence is not entirely correct. The recording of abnormal blood loss takes place after the examination and the consultation. This is then noted, but the patient's complaint was already present before the appointment. Otherwise, you wouldn't go to the doctor.	That's why you go to the hospital; these are (activities) reasons for an intervention, not for a re-intervention.
Did you expect this (after showing chi square significance)?	No	No	Yes
For the question above, why?	The first activity is unclear because a variety of factors have been crammed into it.	It is to be expected that the complaint (abnormal blood loss) is related to the intervention, but not the pattern.	We know from the literature that 10% of people continue to have the complaint and return for a re-intervention.

5.2.2 Local Process Model

For the interview a selection is made to discuss during the interview. For the Local Process Model the selection can be found in Table 5.12. The significant process patterns are highlighted in green based by the chi-square test.

In Table 5.13 the results of the interview questions with the medical expert can be found.

Table 5.12: Selected process patterns of the Local Process Model, which will be discussed during the interview.

Selected process patterns	Support	Confidence	Language fit	Determinism	Coverage
LPM-M1	13	0,70	1	0,66	0,0015
LPM-NM1	626	0,585	1	1	0,0878
LPM-NM2	55	0,196	1	1	0,034

Table 5.13: Qualitative analysis for the Local Process Model.

Questions	LPM-M3	LPM-NM1	LPM-NM2
Is this a process pattern you would expect to be important for the outcome according to your knowledge?	No	No	Yes
For the question above, how confident are you? (Scale from 1-10)	4	8	8
For the question above, why?	Unclear, all those activities seem crammed in. This should be checked with a gynecologist.	This is circular reasoning; you expect that 10% will return and the pattern does not contribute to a re-intervention, but with the reasoning that an intervention does not help 10% of patients, it makes sense that this pattern emerges.	Sometimes you have patients with a less effective coping strategy, who have more difficulty dealing with complaints and thus have more trouble than others. This results in more visits, which could explain the pattern; this could explain a re-intervention.
Did you expect this (after showing chi square significance)?	No	Yes	Yes
For the question above, why?	The terms are unclear.	This is circular reasoning; you expect that 10% will return, and the pattern does not contribute to a re-intervention, but with the reasoning that an intervention does not help 10% of patients, it makes sense that this pattern emerges.	This is due to the coping strategy of the patient and the patient's environment. As a result, you can expect that the patient will keep returning to the clinic. This indicates how a patient deals with the complaint/- does not have a supportive environment and therefore needs a re-intervention more quickly.

Table 5.13 shows that the significant process pattern, LPM-NM2, has an interesting explanation. This process pattern consists of a loop, which aligns with the medical expert's knowledge, indicating its importance for a re-intervention. According to the medical expert, some patients have less effective coping mechanisms, possibly due to a less supportive environment or a lower pain threshold. This results in these patients not managing their complaints well [20]. This could lead to more visits to the clinic and potentially resulting in a re-intervention.

For LPM-M3, one of the events was unclear to the medical expert due to multiple factors being crammed into a single event.

Regarding LPM-NM1, the medical expert did not consider this process pattern important for the outcome, due to circular reasoning. While literature indicates that 10% of patients require a re-intervention, the process pattern itself does not contribute to this outcome.

5.2.3 IMPresseD Framework

For the interview a selection is made to discuss during the interview. For the IMPresseD Framework the selection can be found in Table 5.14. The significant process patterns are highlighted in green based by the chi-square test.

Table 5.14: Selected process patterns of the IMPresseD Framework, which will be discussed during the interview.

Selected process patterns	Frequency interest function	Outcome interest function	Correlation interest function
IMP-M1	0,003	0,002	0,879
IMP-NM1	0,179	0,270	0,820
IMP-NM5	0,031	0,293	0,848

In Table 5.15 the results of the interview questions with the medical expert can be found. As mentioned before all three process patterns are considered significant according to the chi-square tests. For process pattern IMP-M1, the answer is similar to the circular reasoning mentioned before. However, it is not expected to be important according to the medical expert even though it is significant according to the chi-square.

Process pattern IMP-NM1 has a similar answer related to the coping mechanism of the patients. Which results to a higher chance of a re-intervention. Furthermore, the medical expert answers both closed questions with a "Yes"

due to the coping mechanism reasoning. In addition it got the highest confidence score with regards to being important for the outcome.

Process pattern IMP-NM5 is expected to be important for the outcome because everyone receives a consultation after an intervention. However, a consultation is not a predictive factor for re-intervention. Hence why the medical expert gave it a confidence score of seven. What could be interesting is, what is discussed in the consultation. This could be an indication for re-intervention because during a consultation patients can mention that they are unsatisfied with the results, which could result in an re-intervention.

Table 5.15: Qualitative analysis for the IMPresseD Framework.

Questions	**IMP-M1**	**IMP-NM1**	**IMP-NM5**
Is this a process pattern you would expect to be important for the outcome according to your knowledge?	No	Yes	Yes
For the question above, how confident are you? (Scale from 1-10)	9	8	7
For the question above, why?	Causally no, but logically yes, because 10% come back.	Yes, actually the same as a pattern (repetition).	Everyone receives a standard consultation after an intervention, so in that respect the consultation is not predictive of a re-intervention, but the content of the consultation could be decisive for a re-intervention.
Did you expect this (after showing chi square significance)?	Yes	Yes	No
For the question above, why?	It is logical, due to that 10% who return. They come with a complaint, and you know in advance that it won't help 10% of the time.	We have the experience in the clinic that patients have certain characteristics or coping strategies, which indicates how often someone returns.	Because someone says they are not satisfied and therefore still needs a re-intervention. It really depends on the content of the consultations.

5.2.4 Discussion

For the Episode Discovery none of the three process patterns are expected to be important according to the medical expert due to how doctor records

these events in a different order in the system. However, after showing the chi-square score for ED-NM3, the medical expert expected this process pattern to be significant due to 10% of patients coming back for a re-intervention.

The Local Process Model identifies one significant process pattern, which consist of a loop. This process pattern aligns with the medical experts knowledge and indicates its importance for re-intervention. Patients with less effective coping mechanisms may struggle with complaints, leading to more clinic visits and possible re-intervention. According to the medical expert, some patients have less effective coping mechanisms, leading to more clinic visits and possible re-intervention.

For the IMPresseD Framework two process patterns are considered important for the outcome, namely because of comping strategies of patients and because of the fact that everyone receives a standard consultation. However, the latter, is not an indication for re-intervention according to the medical expert.

Based on the qualitative analysis the Local Process Model and the IMPresseD Framework both equally good in discovering process patterns that are aligned with the medical experts expectations. This is due to the fact that they both use multiple quality criteria or interest functions to discover process patterns.

5.3 Quantitative Analysis

5.3.1 Episode Discovery

The results of the quantitative analysis for the Episode Discovery with the logistic regression as a machine learning model, can be found in Table 5.16. Furthermore, the results for the decision tree as a machine learning model can be found in Table 5.17. The code for the quantitative analysis for both the logistic regression model and the decision tree model in Jupyter Notebook can be found in D.

Table 5.16: Quantitative analysis for the Episode Discovery with a logistic regression model, within brackets is the standard deviation.

Episode Discovery - logistic regression	**Mean accuracy**	**Mean F1 Score**	**Mean AUC-ROC value**
Process pattern excluded	*0,871 (±0,027)*	*0,811 (±0,039)*	*0,782 (±0,046)*
ED-M1	0,871 (±0,027)	0,811 (±0,039)	0,789 (±0,043)
ED-M2	0,871 (±0,027)	0,811 (±0,039)	0,783 (±0,045)
ED-M3	0,871 (±0,027)	0,811 (±0,039)	0,798 (±0,046)
ED-NM1	0,871 (±0,027)	0,811 (±0,039)	0,787 (±0,049)
ED-NM2	0,871 (±0,027)	0,811 (±0,039)	0,787 (±0,051)
ED-NM3	0,871 (±0,027)	0,811 (±0,039)	0,798 (±0,048)
ED-NM4	0,871 (±0,027)	0,811 (±0,039)	0,811 (±0,046)
ED-NM5	0,871 (±0,027)	0,811 (±0,039)	0,793 (±0,046)
Average	0,871	0,811	0,792

Table 5.17: Quantitative analysis for the Episode Discovery with a decision tree model, within brackets is the standard deviation.

Episode Discovery - decision tree	**Mean accuracy**	**Mean F1 Score**	**Mean AUC-ROC value**
Process pattern excluded	*0,722 (±0,061)*	*0,765 (±0,050)*	*0,739 (±0,074)*
ED-M1	0,720 (±0,068)	0,763 (±0,056)	0,731 (±0,082)
ED-M2	0,733 (±0,053)	0,773 (±0,046)	0,740 (±0,081)
ED-M3	0,730 (±0,065)	0,771 (±0,055)	0,735 (±0,086)
ED-NM1	0,722 (±0,056)	0,764 (±0,046)	0,720 (±0,077)
ED-NM2	0,716 (±0,065)	0,760 (±0,051)	0,730 (±0,070)
ED-NM3	0,724 (±0,059)	0,766 (±0,048)	0,737 (±0,096)
ED-NM4	0,724 (±0,056)	0,767 (±0,046)	0,724 (±0,065)
ED-NM5	0,714 (±0,061)	0,759 (±0,050)	0,732 (±0,085)
Average	0,723	0,765	0,732

In both tables, the logistic regression and decision tree models do not show any significant outcomes due to the standard deviations. This results in outcomes that are similar with the numbers for the dataset where the process patterns were excluded.

5.3.2 Local Process Model

The results of the quantitative analysis for the Local Process Model with the logistic regression as a machine learning model, can be found in Table 5.18. Furthermore, the results for the decision tree as a machine learning model can be found in Table 5.19. The code for the quantitative analysis for both the logistic regression model and the decision tree model in Jupyter Notebook can be found in D.

Table 5.18: Quantitative analysis for the Local Process Model with a logistic regression model, within brackets is the standard deviation.

Local Process Model - logistic regression	**Mean accuracy**	**Mean F1 Score**	**Mean AUC-ROC value**
Process pattern excluded	*0,871 (±0,027)*	*0,811 (±0,039)*	*0,782 (±0,046)*
LPM-M1	0,871 (±0,027)	0,811 (±0,039)	0,782 (±0,046)
LPM-M2	0,871 (±0,027)	0,811 (±0,039)	0,782 (±0,048)
LPM-M3	0,871 (±0,027)	0,811 (±0,039)	0,790 (±0,055)
LPM-NM1	0,871 (±0,027)	0,811 (±0,039)	0,782 (±0,046)
LPM-NM2	0,871 (±0,027)	0,811 (±0,039)	0,789 (±0,043)
LPM-NM3	0,871 (±0,027)	0,811 (±0,039)	0,790 (±0,049)
LPM-NM4	0,870 (±0,027)	0,811 (±0,039)	0,787 (±0,047)
LPM-NM5	0,871 (±0,027)	0,811 (±0,039)	0,782 (±0,046)
Average	0,871	0,811	0,785

Table 5.19: Quantitative analysis for the Local Process Model with a decision tree model, within brackets is the standard deviation.

Local Process Model - decision tree	**Mean accuracy**	**Mean F1 Score**	**Mean AUC-ROC value**
Process pattern excluded	*0,722 (±0,061)*	*0,765 (±0,050)*	*0,739 (±0,074)*
LPM-M1	0,723 (±0,063)	0,766 (±0,051)	0,729 (±0,081)
LPM-M2	0,718 (±0,062)	0,762 (±0,050)	0,738 (±0,074)
LPM-M3	0,720 (±0,061)	0,764 (±0,050)	0,742 (±0,074)
LPM-NM1	0,723 (±0,060)	0,766 (±0,049)	0,740 (±0,074)
LPM-NM2	0,712 (±0,060)	0,757 (±0,049)	0,734 (±0,064)
LPM-NM3	0,714 (±0,057)	0,758 (±0,045)	0,733 (±0,053)
LPM-NM4	0,722 (±0,061)	0,765 (±0,050)	0,739 (±0,074)
LPM-NM5	0,721 (±0,064)	0,765 (±0,052)	0,739 (±0,077)
Average	0,719	0,763	0,737

Similar to the previous tables, both the logistic regression and decision tree models for the Local Process Model do not show any significant outcomes due to the standard deviations. This results in outcomes that are similar with the numbers for the dataset where the process patterns were excluded.

5.3.3 IMPresseD Framework

The results of the quantitative analysis for the IMPresseD Framework with the logistic regression as a machine learning model, can be found in Table 5.20. Furthermore, the results for the decision tree as a machine learning model can be found in Table 5.21. The code for the quantitative analysis for both the logistic regression model and the decision tree model in Jupyter Notebook can be found in D.

Table 5.20: Quantitative analysis for the IMPresseD Framework with a logistic regression model, within brackets is the standard deviation.

IMPresseD Framework - logistic regression	**Mean accuracy**	**Mean F1 Score**	**Mean AUC-ROC value**
Process pattern excluded	*0,871 (±0,027)*	*0,811 (±0,039)*	*0,782 (±0,046)*
IMP-M1	0,871 (±0,027)	0,811 (±0,039)	0,803 (±0,063)
IMP-M2	0,871 (±0,027)	0,811 (±0,039)	0,781 (±0,047)
IMP-M3	0,871 (±0,027)	0,811 (±0,039)	0,782 (±0,046)
IMP-NM1	0,871 (±0,027)	0,811 (±0,039)	0,789 (±0,054)
IMP-NM2	0,871 (±0,027)	0,811 (±0,039)	0,782 (±0,046)
IMP-NM3	0,870 (±0,028)	0,811 (±0,039)	0,798 (±0,055)
IMP-NM4	0,871 (±0,027)	0,811 (±0,039)	0,783 (±0,046)
IMP-NM5	0,881 (±0,027)	0,846 (±0,037)	0,824 (±0,070)
Average	0,872	0,815	0,792

Table 5.21: Quantitative analysis for the IMPresseD Framework with a decision tree model, within brackets is the standard deviation.

IMPresseD Framework - decision tree	**Mean accuracy**	**Mean F1 Score**	**Mean AUC-ROC value**
Process pattern excluded	*0,722 (±0,061)*	*0,765 (±0,050)*	*0,739 (±0,074)*
IMP-M1	0,723 (±0,061)	0,766 (±0,050)	0,741 (±0,073)
IMP-M2	0,721 (±0,068)	0,764 (±0,056)	0,735 (±0,081)
IMP-M3	0,722 (±0,061)	0,765 (±0,050)	0,736 (±0,075)
IMP-NM1	0,731 (±0,058)	0,771 (±0,048)	0,733 (±0,082)
IMP-NM2	0,722 (±0,061)	0,765 (±0,050)	0,739 (±0,074)
IMP-NM3	0,720 (±0,062)	0,764 (±0,051)	0,737 (±0,074)
IMP-NM4	0,722 (±0,061)	0,765 (±0,050)	0,739 (±0,074)
IMP-NM5	0,723 (±0,064)	0,766 (±0,052)	0,737 (±0,076)
Average	0,722	0,766	0,739

Similar to the previous tables, both the logistic regression and decision tree models for the IMPresseD Framework do not show any significant outcomes due to the standard deviations. This results in outcomes that are similar with the numbers for the dataset where the process patterns were excluded.

5.3.4 Feature Importance

In this section, the results of the feature importance analysis will be discussed. As mentioned in Chapter 4, three analyses were performed. The first analysis, which includes the top 10 events, can be viewed in Figure 5.8 and in Table 5.22. The second analysis incorporates both the top 10 events and the 24 process patterns, as shown in Figure 5.9 and in Table 5.23. Lastly, the third analysis focuses solely on the 24 process patterns, which can be

viewed in Figure 5.10 and in Table 5.24.

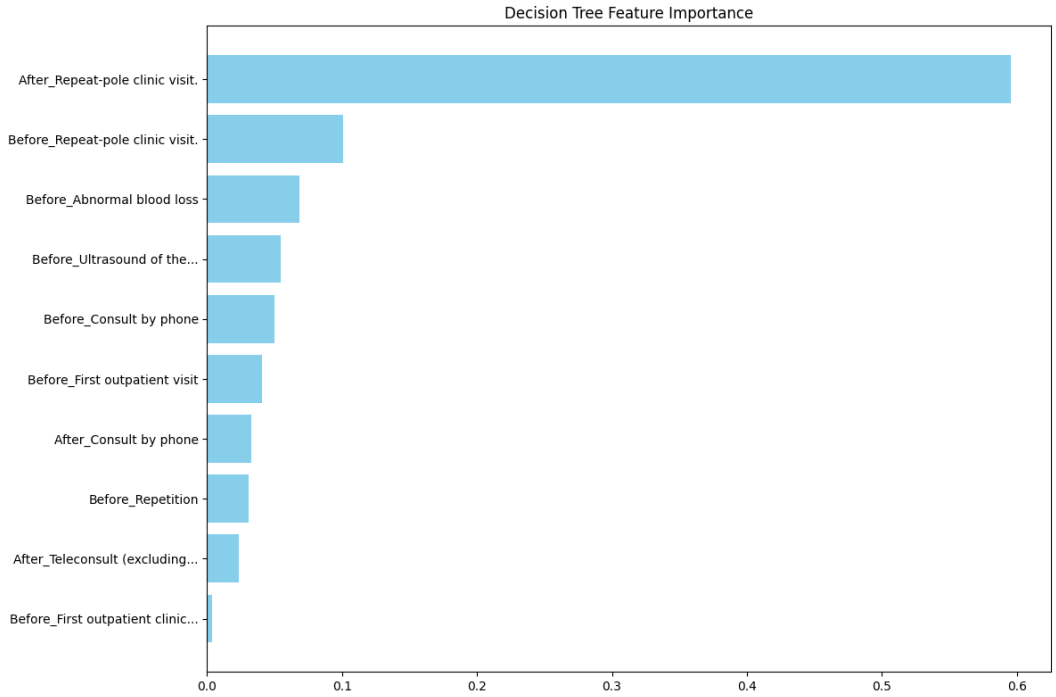

Figure 5.8: Feature importance analysis for the top 10 events which were extracted from Disco.

Table 5.22: Feature importance scores for each feature for the top 10 events, which were extracted from Disco.

Feature	**Importance score**
After_Repeat-pole clinic visit.	0,595494
Before_Repeat-pole clinic visit.	0,100640
Before_Abnormal blood loss	0,068373
Before_Ultrasound of the abdominal organs.	0,054602
Before_Consult by phone	0,049858
Before_First outpatient visit	0,040632
After_Consult by phone	0,032693
Before_Repetition	0,030492
After_Teleconsult (excluding screen to screen image contact, see 190019).	0.023506
Before_First outpatient clinic visit.	0,003711

Based on Figure 5.8 and Table 5.22, the event "After_Repeat-pole clinic visit" has the highest importance score of 0,5954945 among the top 10 events. This indicates that patients who repeatedly visit the clinic after an intervention have a higher likelihood of needing a potential re-intervention. This finding is logical, as mentioned in the interview: patients who frequently visit the clinic likely still have complaints and therefore may require a re-intervention.

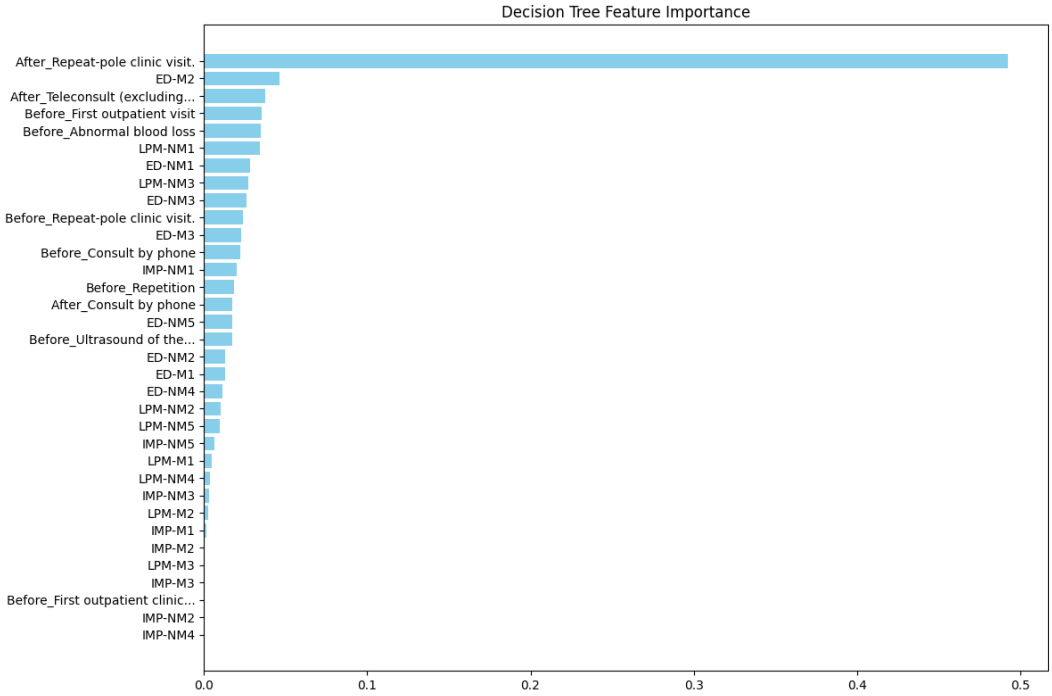

Figure 5.9: Feature importance analysis for the top 10 events and the 24 process patterns.

Table 5.23: Feature importance scores for each feature for the top 10 events and the 24 process patterns.

Feature	Importance score
After_Repeat-pole clinic visit.	0,492409
ED-M2	0,045979
After_Teleconsult (excluding screen to screen image contact, see 190019).	0.037689
Before_First outpatient visit	0,035284
Before_Abnormal blood loss	0,034940
LPM-NM1	0,034359
ED-NM1	0,028425
LPM-NM3	0,027114
ED-NM3	0,026244
Before_Repeat-pole clinic visit.	0,023880
ED-M3	0,022621
Before_Consult by phone	0,022305
IMP-NM1	0,019779
Before_Repetition	0,018170
After_Consult by phone	0,017247
ED-NM5	0,017178
Before_Ultrasound of the abdominal organs.	0,017095
ED-NM2	0,013035
ED-M1	0,012685
ED-NM4	0,011245
LPM-NM2	0,010137
LPM-NM5	0,009783
IMP-NM5	0,006460
LPM-M1	0,004842
LPM-NM4	0,003365
IMP-NM3	0,003237
LPM-M2	0,002752
IMP-M1	0,001381
IMP-M2	0,000362
LPM-M3	0,000000
IMP-M3	0,000000
Before_First outpatient clinic visit.	0,000000
IMP-NM2	0,000000
IMP-NM4	0,000000

In Figure 5.9 and Table 5.23, the event "After_Repeat-pole clinic visit" again stands out with the highest importance score of 0.492409 among the top 10 events and 24 process patterns. This is followed by the medical process pattern ED-M2 from the Episode Discovery method, which includes the events "Before_Sis Echo," "Before_Ultrasound of the abdominal organs," and "Before_Abnormal blood loss." However, as noted during the interview, the order of these events from a patient perspective is incorrect. Patients typically experience a complaint, such as abnormal blood loss, before undergoing medical procedures like an ultrasound and a sis echo. Additionally, the discrepancy between the importance score of "After_Repeat-pole clinic visit" event and the process pattern ED-M2 remains a lot higher.

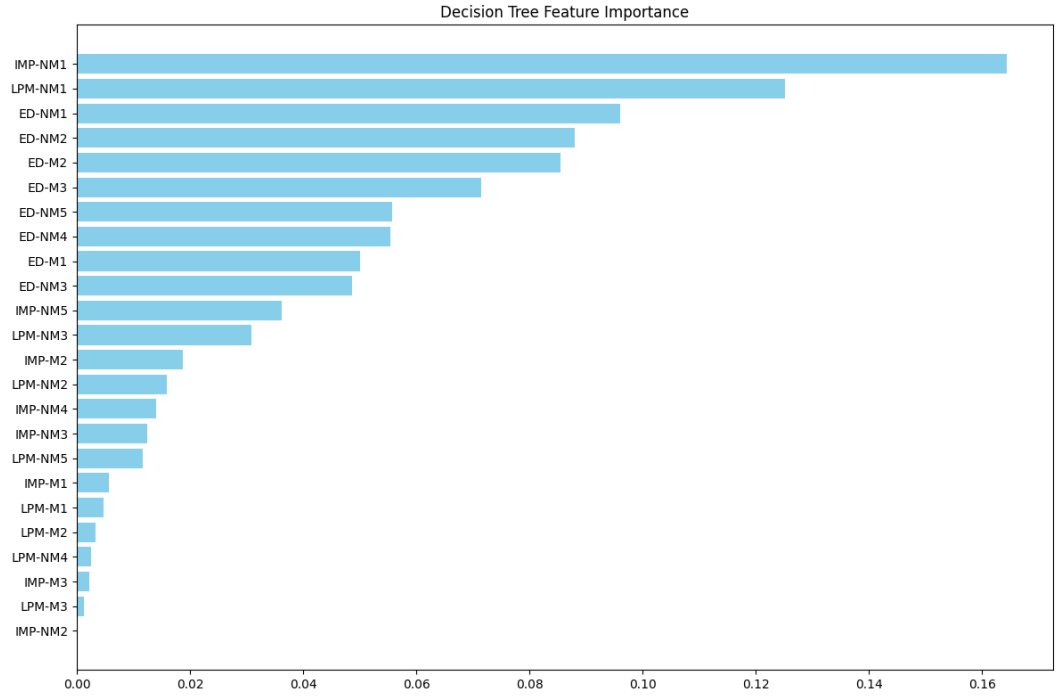

Figure 5.10: Feature importance analysis for the 24 process patterns.

Table 5.24: Feature importance scores for each feature for the 24 process patterns.

Feature	Importance score
IMP-NM1	0,164335
LPM-NM1	0,125114
ED-NM1	0,095963
ED-NM2	0,087937
ED-M2	0,085551
ED-M3	0,071554
ED-NM5	0,055810
ED-NM4	0,055469
ED-M1	0,049999
ED-NM3	0,048602
IMP-NM5	0,036274
LPM-NM3	0,030885
IMP-M2	0,018812
LPM-NM2	0,015835
IMP-NM4	0,013979
IMP-NM3	0,012371
LPM-NM5	0,011662
IMP-M1	0,005710
LPM-M1	0,004694
LPM-M2	0,003383
LPM-NM4	0,002585
IMP-M3	0,002239
LPM-M3	0,001238
IMP-NM2	0,000000

In Figure 5.10 and Table 5.24, the process pattern IMP-NM1 has the highest importance score of 0.164335, which is considerably lower than the score for the event "After_Repeat-pole clinic visit." Despite this, IMP-NM1 includes the event "After_Repeat-pole clinic visit" and is significant according to the chi-square test. The qualitative analysis with the medical expert confirmed that this finding aligns with their knowledge: patients with less effective coping strategies are more likely to need a re-intervention.

Additionally, this figure and table show that LPM-NM1 ranks second in terms of importance score. This process pattern was also discussed during the qualitative analysis. According to literature, medical staff know that 10% of patients return for a re-intervention.

As before, the process pattern ED-NM1 ranks third. However, from a patient perspective, the order of events in this pattern is incorrect due to the way medical staff record information in the system.

5.3.5 Discussion

The logistic regression and decision tree models for the three PPD methods do not show significant outcomes due to high standard deviations. As a result, the outcomes are similar to those from the dataset where the process patterns were excluded.

However, the feature importance analysis reveals that the event "After_-Repeat-pole clinic visit" has the highest importance score among the top 10 events. This makes sense, as discussed in the interview: patients who frequently visit the clinic likely still have complaints and may require re-intervention. Following this is the medical process pattern ED-M2 from the Episode Discovery method. However, as noted during the interview, the order of events in ED-M2 is incorrect from the patient's perspective.

Among the 24 process patterns, IMP-NM1 has the highest importance score, although it is considerably lower than the score for "After_Repeat-pole clinic visit." This is followed by the process patterns LPM-NM1 and ED-NM1. However, as noted during the interview, the order of events in ED-NM1 is incorrect from the patient's perspective, making it less relevant. Similarly, LPM-NM1 is deemed unimportant by the medical expert due to circular reasoning.

Overall, the IMPresseD Framework identifies process patterns that perform better than those from the other two PPD methods. Notably, the process pattern containing "After_Repeat-pole clinic visit," which has the highest feature importance score, is IMP-NM1.

5.4 Representation of PPD Methods

in Table 5.25 the answers of the medical expert are presented with regards to the representation of the models of the three PPD methods.

Table 5.25: Qualitative analysis for the representation of the tree PPD methods.

Questions about representation	**Answers**
Below you see three traces of three patients with their models. Which representation do you prefer?	Model one
For the question above, why?	It's unclear when you get **'OR'**/ **'AND'** in Model one; I interpret it as **'OR'** and not that both **'OR'**/ **'AND'** are possible. In Model three, you cannot deduce the order; it is very vague. Furthermore, I cannot deduce Patient two from Model two. For Model one, I would prefer to see that for Patient one, the two circles are stacked vertically, with one arrow pointing to them.

The responses indicate that the medical expert had a slight preference for model one, which represents the IMPresseD Framework. However, the symbols used in all three PPD methods were unclear to the medical expert. Specifically, for the IMPresseD Framework, it was not clear that both OR and AND gates were possible; the medical expert interpreted it as representing only an OR gateway.

For the Local Process Model, model two, one of the traces in the event log could not be deduced from the local process model according to the medical expert. The symbols related to tokens were particularly unclear to the medical expert.

Regarding model three, which represents the Episode Discovery, the medical expert found the more simplistic representation to be vague and difficult for extracting all the traces from the event log.

Furthermore, table 5.26 provides an overview of the advantages and dis-

Table 5.26: GUI capabilities for each PPD method.

GUI	ED	LPM	IMP
Ability to extend upon chosen process pattern	No	No	Yes
Ability to select outcome column	No	No	Yes
Ability to exclude/include events*	Yes**	No	No
Ability to order the results from the discovery	No	Yes	No***

* Episodes does have a function to include only certain events in the process pattern discovery. However, this becomes impractical when there are too many events.

** Ideally, it would be useful to exclude events like "Consult by phone" from the pattern discovery, as they are less medically significant compared to events like "Ultrasound of abdominal organs."

*** For the IMPResseD Framework, ordering each interest function from largest to smallest, and vice versa, would be beneficial. This approach allows for faster filtering.

advantages of each PPD method. Throughout this research, the three PPD methods have been extensively used for the systematic selection. As a result, a comprehensive overview of the methods' graphical user interfaces (GUIs) can be presented.

Based on the GUI capabilities outlined in Table 5.26, the IMPresseD Framework offers the most features. In particular, the ability to select the outcome column is especially useful, as it ensures that the correct column is considered for the outcome interest function. Furthermore, it has the ability to extend upon a chosen pattern, making sure that extension are possible and thus further analysis.

In addition, the Episode Discovery is the only PPD method in this research that has the ability to exclude events. Especially for this dataset, it would have been useful to exclude events that were, medically speaking less interesting for the other two PPD methods as well.

Furthermore, the Local Process Model has the ability to order results from the process pattern discovery from smallest to largest. Which comes in handy when selecting process patterns based on certain quality criteria. For the IMPresseD Framework this would have been useful as well, for example to order the results based on a interest function. For the Episode Discovery this is automatically done, because the PPD method only focuses on frequency of the discovered process patterns.

6. Discussion

In this research, various aspects of Episode Discovery, Local Process Models, and the IMPResseD Framework were explored and compared within a medical setting. Although the findings provide valuable insights, several limitations must be acknowledged to fully understand the study's scope and implications. Additionally, several areas for future work are recommended to advance the field further.

This variance between the discovered process patterns of the three different PPD methods, might be related to the algorithms used to discover process patterns. However, the interview will serve as an alignment check to verify if these are the results a medical expert would expect based on their expertise and knowledge.

Based on systematic selection, qualitative, and quantitative analysis, the IMPresseD Framework performs the strongest in a medical setting among the three PPD methods. It identified three significant process patterns in the chi-square test. In the qualitative analysis, two of these three patterns were deemed important by the medical expert for the outcome. Additionally, in the quantitative analysis, the process pattern IMP-NM1 ranked among the highest in feature importance. This strong performance may be attributed to the interest functions used by the IMPresseD Framework, particularly the outcome interest function, which assesses how much a discovered process pattern contributes to the outcome—in this case, re-intervention.

The Local Process Model also performs well, identifying one significant process pattern in the systematic selection. This pattern was considered particularly interesting in the qualitative analysis due to its loop, which relates to patients' coping strategies. However, in the quantitative analysis, the process pattern that scored highest was deemed unimportant due to circular reasoning.

Overall, Episode Discovery performed the worst. Although it identified two significant process patterns in the chi-square test, these were significant only because they had low actual frequencies. In the qualitative analysis, the medical expert did not expect any of the three process patterns to be important for the outcome. While one pattern scored highest in feature importance, further investigation revealed that the order of events was incorrect from the patient's perspective. The poor performance of Episode Discovery

may be due to its focus solely on process pattern frequencies, whereas the other two PPD methods consider additional factors in discovering process patterns.

This answers the main research question, namely *How do various process pattern detection methods perform on Novasure data set, and what insights can be gained through this comparison?*, the IMPresseD Framework performs the strongest due to its interest functions and its ability to select a column in the event log that is considered the outcome of the dataset, which is related to the outcome interest function. The Local Process Model also performs quite well due to its five quality criteria, resulting in a more comprehensive discovery of process patterns. Episode Discovery is the least successful of the three on the Novasure dataset because it focuses solely on a single dimension, frequency, making it more singularly focused in discovering process patterns.

6.1 Scientific implications

Many process pattern mining techniques have been developed such as the four PPD methods discussed in the related work in chapter 3. However, more PPD methods exists such as the ARE miner which focuses on action-response-effect patterns [13]. In this paper, the ARE miner is being tested on a healthcare process in a Dutch residential care facility. Furthermore, another technique that is used, is the frequent sequence patterns, which is used on a medical dataset to learn the behavioral criteria of a group of patients. This is then used to cluster groups of patients together based on the rank of the pattern [21].

These PPD methods have been individually tested on medical datasets, but a direct comparison on the same dataset has never been conducted. Such a comparison could be highly beneficial, as it would reveal which PPD method performs optimally on specific medical dataset, thereby closing a significant gap in the scientific community. This gap concerns the lack of comparative studies on various PPD methods, which could otherwise highlight their relative strengths and weaknesses of each PPD method. To address this, the current research used the Novasure dataset to conduct a comprehensive comparative study, evaluating the effectiveness of three different PPD methods in a medical setting.

Furthermore, the insights gained from this study could inspire the devel-

opment of new PPD methods that combine the strengths of existing ones, leading to more robust and versatile tools for process pattern detection in healthcare.

6.2 Practical implications

The study contributed valuable insights to the field of process pattern mining in healthcare, providing practical recommendations for medical experts and researchers on the most effective PPD methods for analyzing healthcare processes. These insights can be directly applied by other professionals in the process pattern mining field, both in academia and industry, to select and implement the most suitable PPD methods for their medical dataset. Therefore, medical staff can select a PPD method and use it to analyze their medical dataset to find process patterns that might related to the outcome of a medical procedure.

Furthermore, the standardized approach used in this study can serve as a reference model for future comparative studies, enabling other researchers to conduct similar comparative evaluations across different domains and with different PPD methods.

6.3 Limitations

While the findings offer valuable insights, it is important to acknowledge several limitations to fully understand the scope and implications of this research.

One limitation is the process pattern discovery within the Episode Discovery. A K parameter is used within the Episode Discovery to ensure that a minimum of K events must occur before they are counted towards the frequency. As a result, the frequency numbers extracted from Disco may differ from those obtained from Episode Discovery, potentially leading to inaccuracies in the chi-square test results. However, this issue has been partially mitigated by selecting the following options in Disco: setting the 'Time between matching events must be' to 'shorter' and specifying '1 day'. Since most events within a process pattern occur within a single day, this configuration helps align the frequency numbers from Disco with those from Episode Discovery more closely.

Additionally, in Disco, some caseIDs might overlap because a process

pattern could appear multiple times within the process of a single caseID. Moreover, there might be instances where the process pattern does not appear at all. This issue can occur when using filters for follow-up 1 and 2 in Disco. If follow-up 1 and 2 are not connected, they may appear separately within a caseID's process, rather than as part of a cohesive process pattern. Consequently, this separation can result in the process pattern not being identified correctly. Which results in extracting the wrong numbers from Disco to calculate the chi-square tests. However, random checks were performed to mitigate this issue, and this issue was observed only with Episode Discovery and Local Process Models. For the IMPresseD Framework, the numbers could be directly extracted from the IMPresseD software itself.

Another limitation and its impacts is the number of process patterns discussed during the interview. Out of the 24 selected process patterns, only 9 were discussed during the interview. This limited discussion means that not all patterns were reviewed or evaluated in detail, potentially leading to an incomplete assessment of the entire set. Focusing on only 9 patterns could impact the overall understanding and differences in the process pattern discovery of the three PPD methods. The reason for this selective discussion was time limitations.

Another point to consider, for the quantitative analysis, all 24 process patterns were used for logistic regression and decision tree models. However, ideally, at least 30 process patterns should be used to make a robust comparison and achieve more reliable results.

Moreover, the distribution was too limited for the model to learn effectively, as only 129 out of 1,000 cases needed an re-intervention, resulting in an imbalance in the qualitative analysis. However, this issue has been addressed using GridSearchCV in the Jupyter notebook. GridSearchCV optimized hyperparameters by systematically evaluating different parameter combinations, which improved the model's performance. This process helped mitigate the imbalance by enhancing the model's ability to generalize across the limited data, thereby improving its predictive accuracy.

6.4 Future Work

During the course of preparing and performing this research, several areas for future work on PPD methods could benefit the scientific community.

One such area is conducting a systematic literature review of all the

PPD methods that exist within the scientific community. Currently, there are many PPD methods known within the scientific community, specifically to the process mining group. However, a systematic literature review has not yet been performed. Hence, conducting such research would be valuable. Conducting a thorough and comprehensive review of existing literature about the various PPD methods would allow researchers to build on the current body of knowledge, identify gaps, contribute meaningful insights, and understand the differences among the various PPD methods, including their advantages and disadvantages. This would lead to a more informed selection of the most suitable PPD method for their specific area of interest and the context of their event log.

Another interesting area of study, specifically for the IMPResseD Framework, is the creation of a custom interest function. Within the IMPResseD Framework, it is possible to develop such a function if one has programming knowledge. For an impressed framework, it would be beneficial to create a custom interest function which integrates the chi-square test. The chi-square test can help determine the statistical significance of the observed differences between expected and actual groups within the event log.

7. Conclusion

In this chapter, the conclusions from this research will be presented based on the results achieved. Specifically, this chapter will address the main research question, *How do various process pattern detection methods perform on Novasure data set, and what insights can be gained through this comparison?*. The conclusion will be structured around answering this main question through its sub-questions.

Based on the results, the process patterns discovered by the Local Process Model and IMPresseD framework are associated with cases that require re-intervention. This indicates that these process pattern discovery methods are effective in identifying patterns within a medical dataset that are related to the outcome of a procedure. Specifically, they extracted process patterns that can explain why a patient might need a re-intervention, such as frequent clinic visits due to ineffective coping mechanisms.

The reason these two PPD methods are more effective than Episode Discovery might be that they consider multiple dimensions rather than focusing solely on frequency, as Episode Discovery does. This could result in more meaningful and interesting discovery of process patterns.

This finding aligns with the qualitative analysis, which suggests that Episode Discovery is less consistent with the medical expert's knowledge. The expert noted that the order of events often does not match the patient's perspective, and some activities in the process patterns are not directly related to the reasons for a re-intervention. This indicates that Episode Discovery is less effective at identifying relevant process patterns in a medical dataset.

Interestingly, the process pattern LPM-NM2, which is significant according to the chi-square test, includes a loop that caught the medical expert's attention. This is because patients have various coping mechanisms for dealing with their complaints, which can lead to recurring visits. The feature importance analysis supports this observation, as the event "After_-Repeat-pole clinic visit" and the process pattern IMP-NM1, which includes this event, received the highest scores. The Local Process Model effectively represents this recurring pattern as a loop. This capability is not present in the Episode Discovery and IMPresseD Framework methods. Therefore, the Local Process Model proves to be a particularly valuable PPD method due

to its ability to represent loops within process patterns, making it especially useful in a medical setting.

However, the representation of these models for medical experts requires improvement. During the qualitative analysis, extracting traces from the process pattern models proved challenging due to the Petri net notation used in the Local Process Model. The various symbols and tokens are difficult for medical experts to interpret. On the other hand, the simpler notation of the Episode Discovery and the IMPresseD Framework is easier to understand but also tends to be ambiguous. For instance, the Episode Discovery represents parallelism by showing events stacked on top of each other, as seen in Figure D.1. However, the medical expert perceives this differently, interpreting it as a choice between one event or another, rather than as parallel events. The IMPresseD Framework faces similar issues due to its simpler notation.

Overall, the Local Process Model and IMPresseD Framework perform effectively in medical settings by discovering process patterns that are relevant to the outcomes of medical procedures. Their effectiveness stems from their use of multiple dimensions in the analysis. Considering dimensions such as frequency and their relation to the outcome leads to a more meaningful discovery of process patterns in a medical context.

Bibliography

[1] M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management*. Springer, 2018, vol. 2.

[2] W. van der Aalst, "Process mining: Overview and opportunities," *ACM Transactions on Management Information Systems*, vol. 3, no. 2, pp. 1–17, Jul. 2012, ISSN: 2158-6578. DOI: 10.1145/2229156.2229157. [Online]. Available: http://dx.doi.org/10.1145/2229156.2229157.

[3] A. Guastalla, E. Sulis, R. Aringhieri, S. Branchi, C. Di Francescomarino, and C. Ghidini, "Workshift scheduling using optimization and process mining techniques: An application in healthcare," in *2022 Winter Simulation Conference (WSC)*, 2022, pp. 1116–1127. DOI: 10.1109/WSC57314.2022.10015405.

[4] E. Rojas, J. Munoz-Gama, M. Sepúlveda, and D. Capurro, "Process mining in healthcare: A literature review," *Journal of Biomedical Informatics*, vol. 61, pp. 224–236, Jun. 2016, ISSN: 1532-0464. DOI: 10.1016/j.jbi.2016.04.007. [Online]. Available: http://dx.doi.org/10.1016/j.jbi.2016.04.007.

[5] M. Leemans and W. M. P. van der Aalst, "Discovery of frequent episodes in event logs," in *Data-Driven Process Discovery and Analysis*. Springer International Publishing, 2015, pp. 1–31, ISBN: 9783319272436. DOI: 10.1007/978-3-319-27243-6_1. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-27243-6_1.

[6] J. Munoz-Gama, N. Martin, C. Fernandez-Llatas, *et al.*, "Process mining for healthcare: Characteristics and challenges," *Journal of Biomedical Informatics*, vol. 127, p. 103 994, 2022.

[7] M. Vazifehdoostirani, L. Genga, X. Lu, R. Verhoeven, H. van Laarhoven, and R. Dijkman, "Interactive multi-interest process pattern discovery," in *Business Process Management*. Springer Nature Switzerland, 2023, pp. 303–319, ISBN: 9783031416200. DOI: 10.1007/978-3-031-41620-0_18. [Online]. Available: http://dx.doi.org/10.1007/978-3-031-41620-0_18.

[8] Z. Huang, X. Lu, and H. Duan, "Latent treatment pattern discovery for clinical processes," *Journal of Medical Systems*, vol. 37, no. 2, Feb. 2013, ISSN: 1573-689X. DOI: 10.1007/s10916-012-9915-2. [Online]. Available: http://dx.doi.org/10.1007/s10916-012-9915-2.

[9] *Maxima medical center novasure*, https://www.mmc.nl/gynaecologie/aandoeningen-en-behandelingen/novasure/, Accessed on the 8th of January 2024.

[10] H. Xie, Y. Wan, S. Yi, *et al.*, "Clinical analysis of 2152 cases of abnormal uterine bleeding treated by novasure endometrial ablation," *International Journal of Gynecology amp; Obstetrics*, vol. 158, no. 2, pp. 301–307, Nov. 2021, ISSN: 1879-3479. DOI: 10.1002/ijgo.13983. [Online]. Available: http://dx.doi.org/10.1002/ijgo.13983.

[11] Y. Levy-Zauberman, A.-G. Pourcelot, P. Capmas, and H. Fernandez, "Update on the management of abnormal uterine bleeding," *Journal of Gynecology Obstetrics and Human Reproduction*, vol. 46, no. 8, pp. 613–622, Oct. 2017, ISSN: 2468-7847. DOI: 10.1016/j.jogoh.2017.07.005. [Online]. Available: http://dx.doi.org/10.1016/j.jogoh.2017.07.005.

[12] R. Gimpelson, "Ten-year literature review of global endometrial ablation with the novasureamp;reg; device," *International Journal of Women's Health*, p. 269, Mar. 2014, ISSN: 1179-1411. DOI: 10.2147/ijwh.s56364. [Online]. Available: http://dx.doi.org/10.2147/IJWH.S56364.

[13] J. J. Koorn, X. Lu, H. Leopold, and H. A. Reijers, "Looking for meaning: Discovering action-response-effect patterns in business processes," in *Business Process Management*. Springer International Publishing, 2020, pp. 167–183, ISBN: 9783030586669. DOI: 10.1007/978-3-030-58666-9_10. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-58666-9_10.

[14] X. Lu, S. A. Tabatabaei, M. Hoogendoorn, and H. A. Reijers, "Trace clustering on very large event data in healthcare using frequent sequence patterns," in *Business Process Management*. Springer International Publishing, 2019, pp. 198–215, ISBN: 9783030266196. DOI: 10.1007/978-3-030-26619-6_14. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-26619-6_14.

[15] N. Tax, N. Sidorova, R. Haakma, and W. M. van der Aalst, "Mining local process models," *Journal of Innovation in Digital Ecosystems*, vol. 3, no. 2, pp. 183–196, Dec. 2016, ISSN: 2352-6645. DOI: 10.1016/j.jides.2016.11.001. [Online]. Available: http://dx.doi.org/10.1016/j.jides.2016.11.001.

[16] R. J. C. Bose and W. M. van der Aalst, "Discovering signature patterns from event logs," in *2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, IEEE, Apr. 2013. DOI: 10.1109/cidm.2013.6597225. [Online]. Available: http://dx.doi.org/10.1109/CIDM.2013.6597225.

[17] F. Mannhardt, M. de Leoni, H. A. Reijers, W. M. v. d. Aalst, and P. J. Toussaint, "Guided process discovery – a pattern-based approach," *Information Systems*, vol. 76, pp. 1–18, Jul. 2018, ISSN: 0306-4379. DOI: 10.1016/j.is.2018.01.009. [Online]. Available: http://dx.doi.org/10.1016/j.is.2018.01.009.

[18] C. J. Turner, A. Tiwari, R. Olaiya, and Y. Xu, "Process mining: From theory to practice," *Business process management journal*, vol. 18, no. 3, pp. 493–512, 2012.

[19] W. M. van der Aalst, S. Guo, and P. Gorissen, "Comparative process mining in education: An approach based on process cubes," in *Data-Driven Process Discovery and Analysis: Third IFIP WG 2.6, 2.12 International Symposium, SIMPDA 2013, Riva del Garda, Italy, August 30, 2013, Revised Selected Papers 3*, Springer, 2015, pp. 110–134.

[20] S. L. Popa, M. Fadgyas Stanculete, S. Grad, *et al.*, "Coping strategies and inflammatory bowel disease: A narrative review," *Journal of Clinical Medicine*, vol. 13, no. 6, 2024, ISSN: 2077-0383. DOI: 10.3390/jcm13061630. [Online]. Available: https://www.mdpi.com/2077-0383/13/6/1630.

[21] X. Lu, S. A. Tabatabaei, M. Hoogendoorn, and H. A. Reijers, "Trace clustering on very large event data in healthcare using frequent sequence patterns," in *Business Process Management*, T. Hildebrandt, B. F. van Dongen, M. Röglinger, and J. Mendling, Eds., Cham: Springer International Publishing, 2019, pp. 198–215, ISBN: 978-3-030-26619-6.

A. Systematic Selection Models

A.1 Episode Discovery

Below the models from the systematic selection, of both medical and non-medical process pattern, from the Episode Discovery are visible:

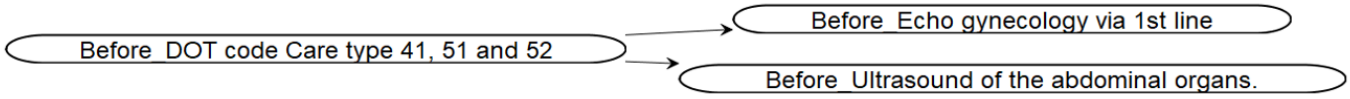

Figure A.1: Process pattern ED-M1 from the Episode Discovery

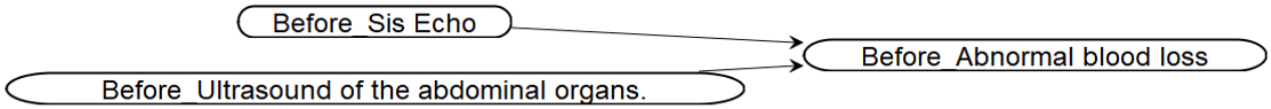

Figure A.2: Process pattern ED-M2 from the Episode Discovery

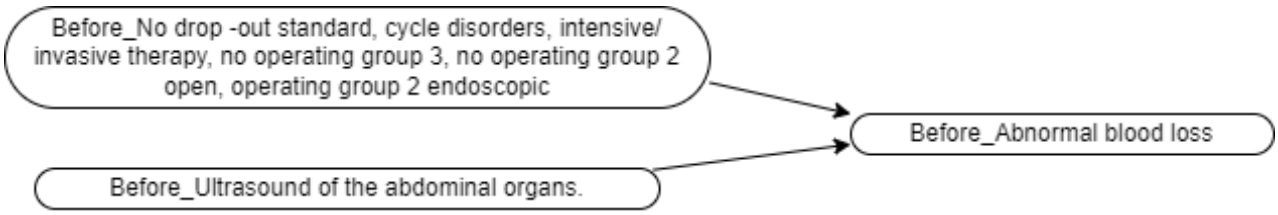

Figure A.3: Process pattern ED-M3 from the Episode Discovery. Due to readability, this process pattern has been remade.

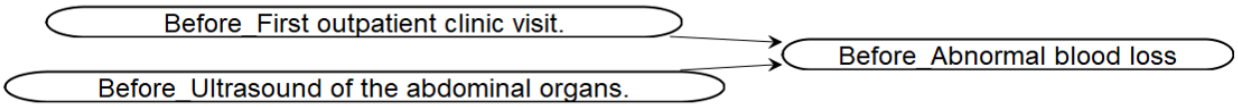

Figure A.4: Process pattern ED-NM1 from the Episode Discovery

Figure A.5: Process pattern ED-NM2 from the Episode Discovery

Figure A.6: Process pattern ED-NM3 from the Episode Discovery

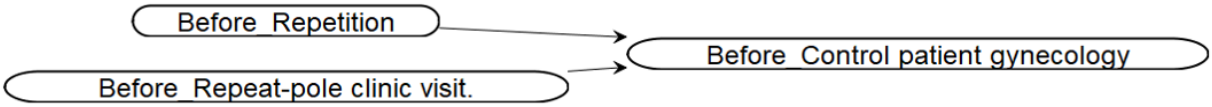

Figure A.7: Process pattern ED-NM4 from the Episode Discovery

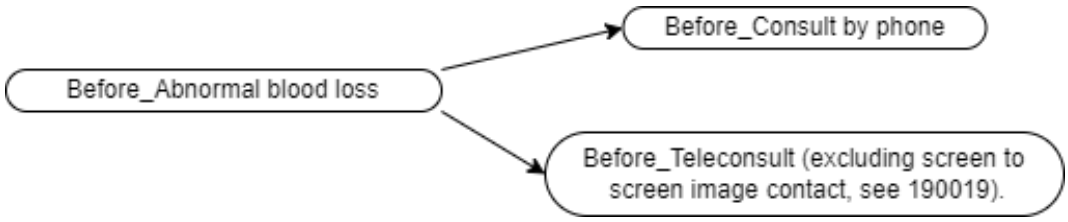

Figure A.8: Process pattern ED-NM5 from the Episode Discovery. Due to readability, this process pattern has been remade.

A.2 Local Process Model

Below the models from the systematic selection, of both medical and non-medical process pattern, from the Local Process Model are visible:

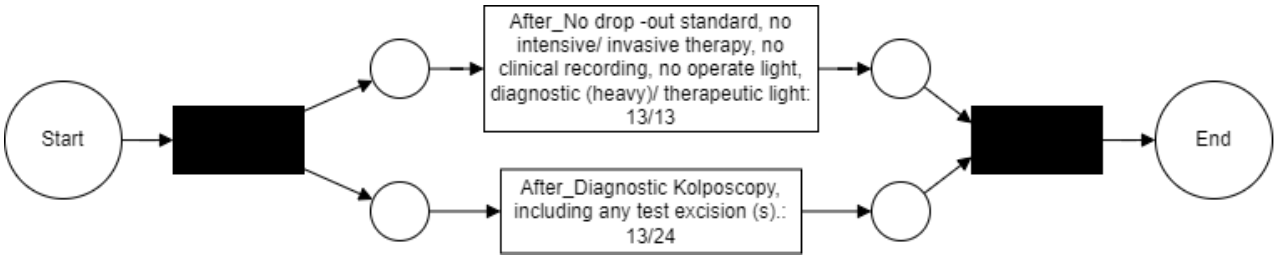

Figure A.9: Process pattern LPM-M1 from the Local Process Model. Due to readability, this process pattern has been remade.

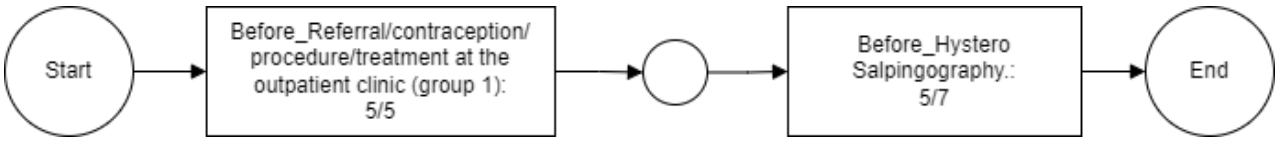

Figure A.10: Process pattern LPM-M2 from the Local Process Model. Due to readability, this process pattern has been remade.

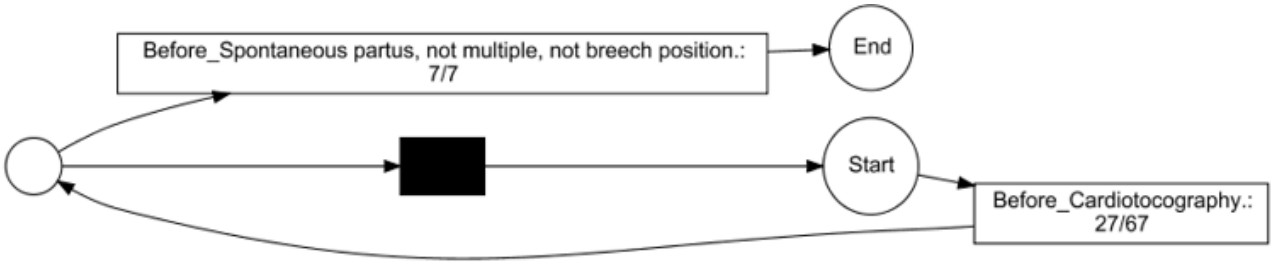

Figure A.11: Process pattern LPM-M3 from the Local Process Model

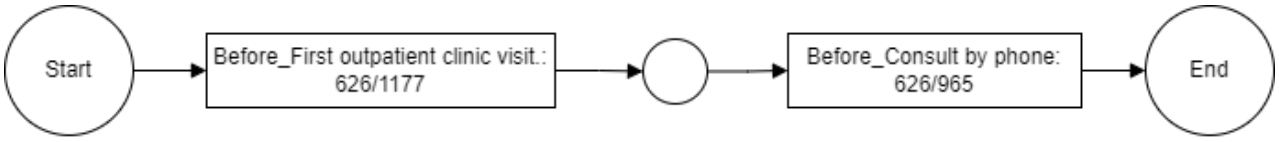

Figure A.12: Process pattern LPM-NM1 from the Local Process Model. Due to readability, this process pattern has been remade.

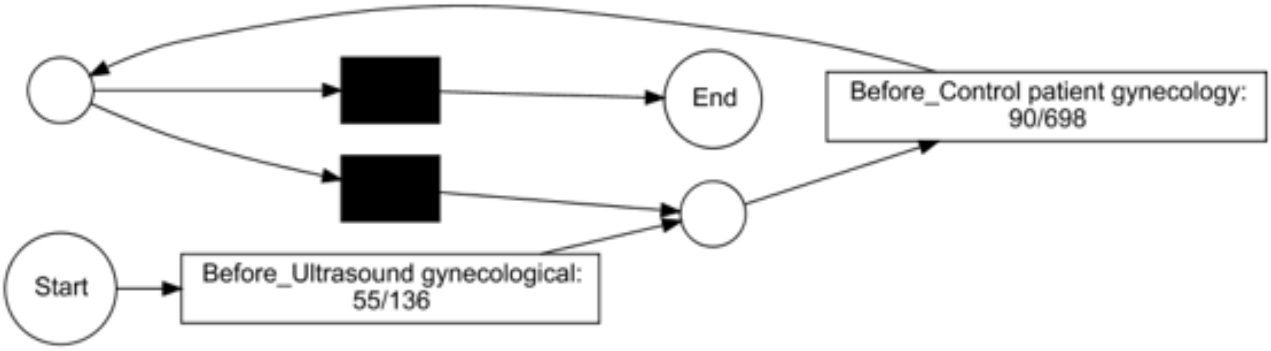

Figure A.13: Process pattern LPM-NM2 from the Local Process Model

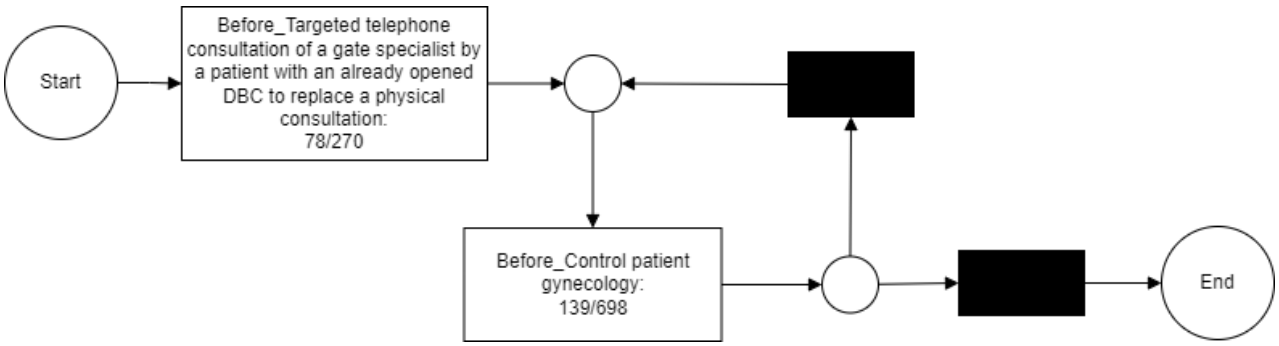

Figure A.14: Process pattern LPM-NM3 from the Local Process Model. Due to readability, this process pattern has been remade.

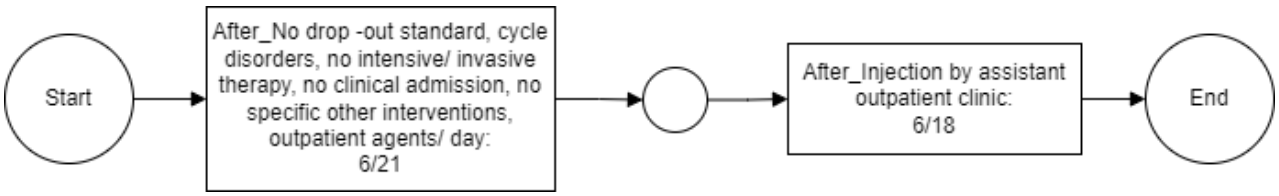

Figure A.15: Process pattern LPM-NM4 from the Local Process Model. Due to readability, this process pattern has been remade.

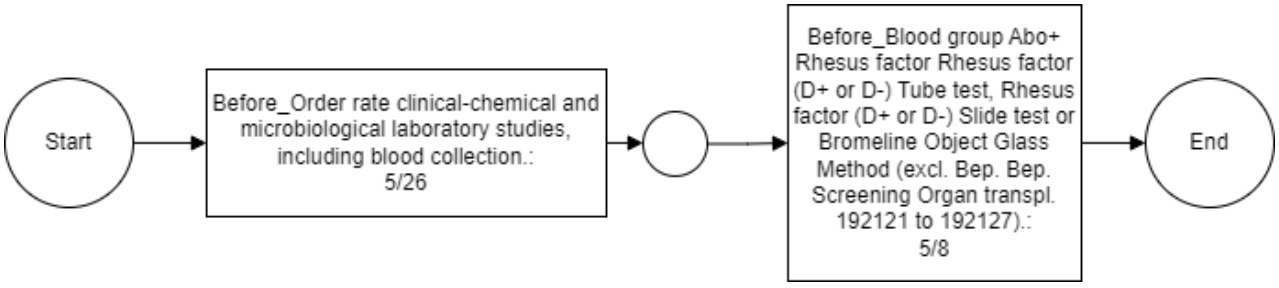

Figure A.16: Process pattern LPM-NM5 from the Local Process Model. Due to readability, this process pattern has been remade.

A.3 IMPresseD Framework

Below the models from the systematic selection, of both medical and non-medical process pattern, from the IMPresseD Framework are visible:

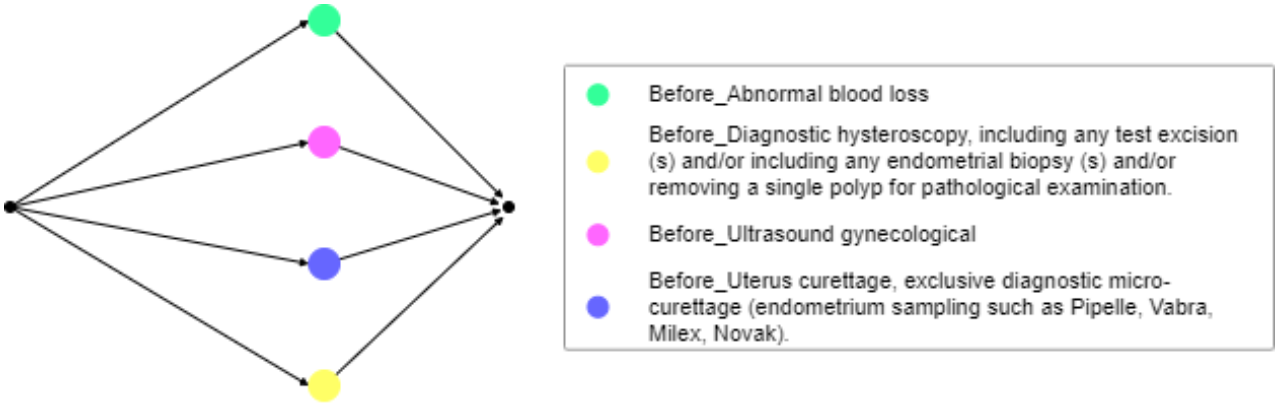

Figure A.17: Process pattern IMP-M1 from the IMPresseD Framework. Due to readability, this process pattern has been remade.

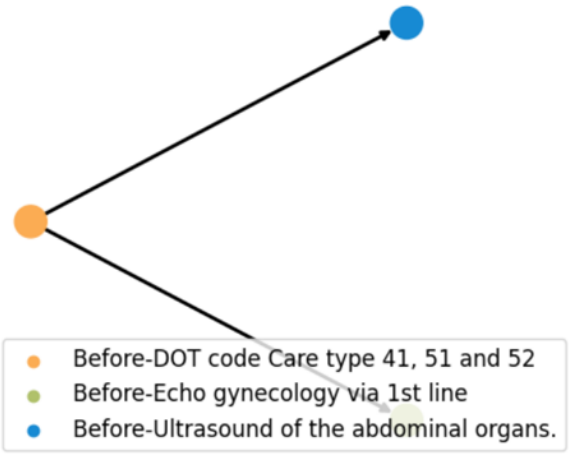

Figure A.18: Process pattern IMP-M2 from the IMPresseD Framework

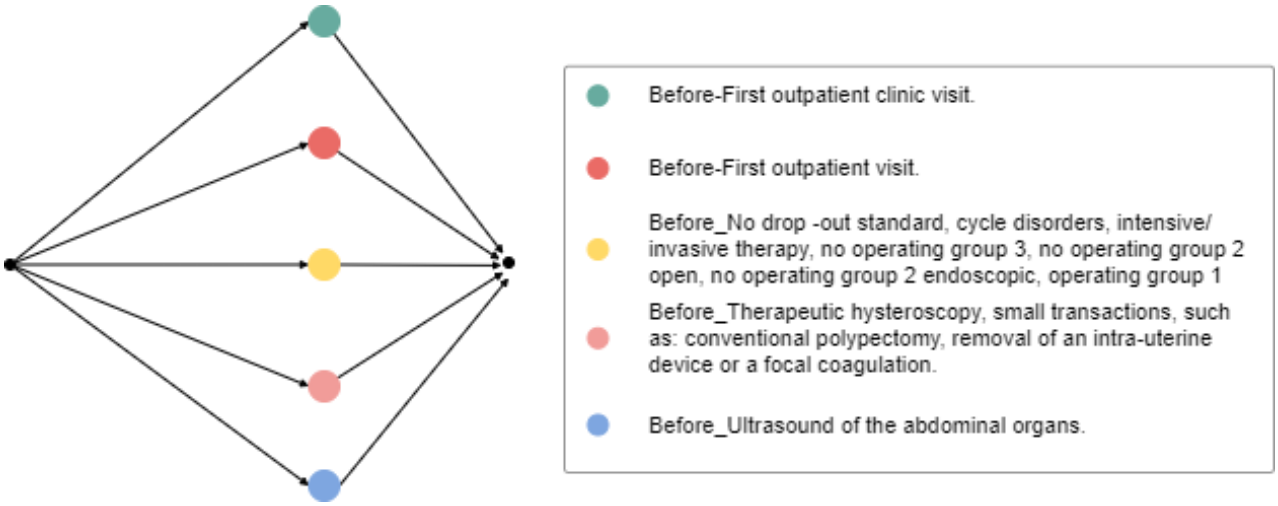

Figure A.19: Process pattern IMP-M3 from the IMPresseD Framework. Due to readability, this process pattern has been remade.

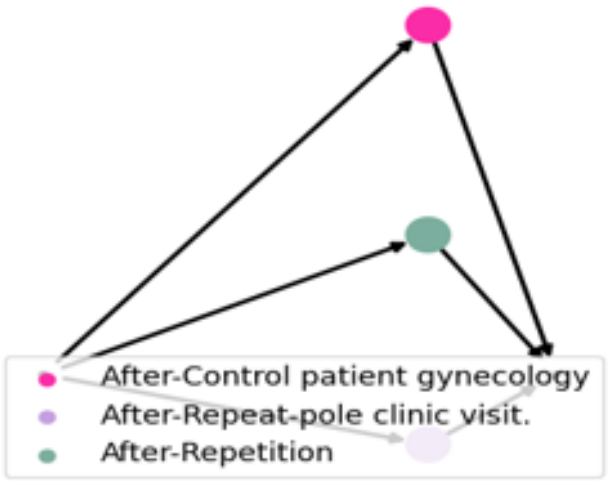

Figure A.20: Process pattern IMP-NM1 from the IMPresseD Framework

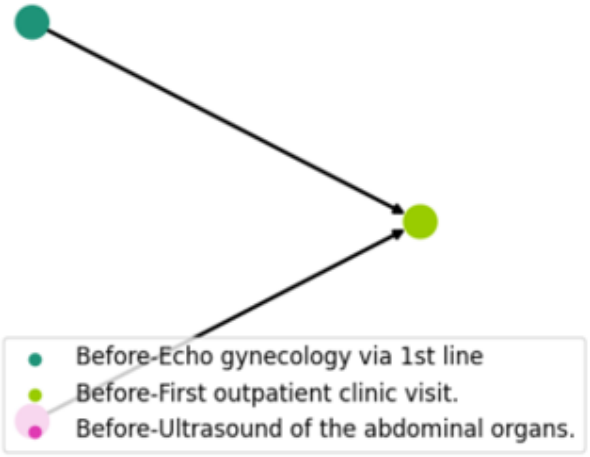

Figure A.21: Process pattern IMP-NM2 from the IMPresseD Framework

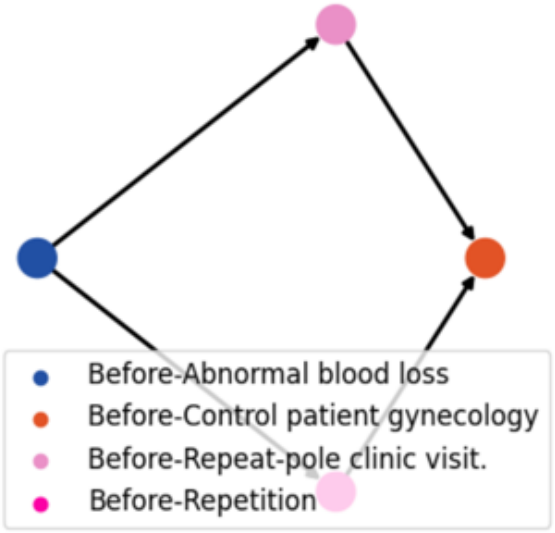

Figure A.22: Process pattern IMP-NM3 from the IMPresseD Framework

Figure A.23: Process pattern IMP-NM4 from the IMPresseD Framework

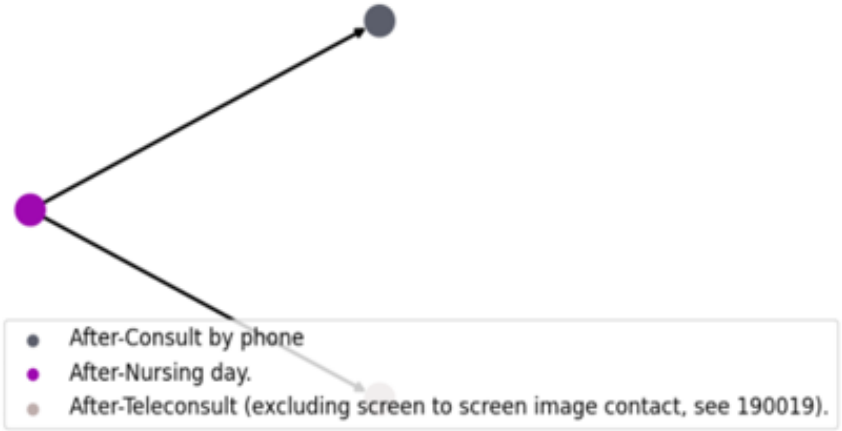

Figure A.24: Process pattern IMP-NM5 from the IMPresseD Framework

B. Interview Transcript

B.1 First Process Pattern of Episode Discovery

Sabrina: Dit is dus eigenlijk een patroon die ik in de dataset heb gevonden . We hebben zeg maar verschillende frameworks, dus verschillende methodes om die patronen te achterhalen en die patronen hebben dan ook een algoritme. En dan zegt een framework van, deze patroon is belangrijk voor de uitkomst van de Novasure Surgery. En dat is in dit geval interventie of geen interventie. En nou ja, het patroon in dit geval is met de Episode Discovery ontdekt, dat is dus een van die methodes. En wat je ziet is dat Before first outpatient clinic visit voor de operatie plaats vindt en dat is een event of activiteit dat plaats vindt. Tegelijk of misschien daarvoor of daarna vindt Before ultrasound of the abdominal organs plaats. Als die twee stappen gebeurd zijn, dan heb je daarna Before_Abnormal blood loss activiteit. Mijn vraag is dan, nouja, het staat hier al op. Is this a process pattern you would expect to be important for the outcome according to your knowledge? En dan gebaseerd op jouw kennis.

Jaklien: Ik moet even nadenken hoor, dus jij zegt van links staat eigenlijk van. Ja, dat is het eerste consult, samen met een ultrasound, maar ik snap alleen niet before abnormal blood loss. Kijk ze komen met abnormaal bloedverlies he? Je moet zo zien, ze hebben een klacht.

Sabrina: Ja.

Jaklien: Abnormaal uterine bloedverlies en daarvoor komen ze naar de kliniek en dan krijgen ze dus een consult en een ultra sound. En dan op basis daarvan gaan ze vaststellen dat ze een novasure bijvoorbeeld gaan doen. Dus de blood loss is als eerste. Met die klacht komen ze en dan hebben ze dus die dat consult en het onderzoek en dan daarna krijgen ze de novasure.

Sabrina: Oke, dus eigenlijk is het wel standaard procedure dat ze eerst met abnormaal bloedverlies komen.

Jaklien: Ja, dus eigenlijk komt die de andere kant, dus dat rechter blok moet links staan. En dan krijg je die twee, die boven elkaar staan en dan moet er rechts staan de novasure.

Sabrina: Ja, maar dit is een patroon binnen het hele proces zelf. Dus die before, dat moet je eigenlijk negeren. Dat was voor de dataset zelf, want ik heb dus verschillende datasets en dat before woordje daarmee bedoel, dat het voor de novasure operatie plaat vindt en je hebt ook patronen waarbij staat after die beginnen dan met after als het na de operatie zelf is. Maar dat was weer om voor mij het onderscheid te maken tussen welke voor en welk na de operatie plaats vindt?

Jaklien: Oke en dan is het dus zo dat je dan de abnormal blood loss, daar wil jij dan mee zeggen: de klacht is niet veSrholpen en daarom krijg je een re-interventie? Is dat de abnormal blood loss, of niet?

Sabrina: Oh, dat weet ik nu even niet. Dit is hoe de activiteit werd genoemd in de dataset, dus echt alleen abnormal blood loss stond erbij.

Jaklien: Nou, kijk, je moet het zo zien, h. Kijk, de uiteindelijke uitkomst maat is dus re-interventie. Dat wil dus zeggen dat de novasure niet voldoende heeft geholpen. Er blijft bloedverlies aanhouden, dus gaan we nog een interventie doen en dat noem je dan een re-interventie. Omdat het eigenlijk een tweede interventie is. Als jij mij dit bedoelt, h? Van je krijgt een visite en een ultrasound en die laat zien dat de klacht onvoldoende geholpen is, want ze hebben nog steeds abnormaal bloedverlies. Dat is, wat je wil zeggen hier of niet?

Sabrina: Nee, wat ik met dit patroon bedoel en wat ik uit de dataset heb gehaald, je hebt zeg maar die dataset, die bestaat uit verschillende activiteiten, dus een klant bezoekt de kliniek. Er wordt een ultrasound gemaakt. De klant heeft een klacht, klant. Sorry, de patint heeft een klacht en dat is abnormaal bloed verlies.

Jaklien: Ja.

Sabrina: Ja, en dat dat is dan heel proces, h? Hoe een patint door het ziekenhuis gaat.

Jaklien: Ja.

Sabrina: En dan hopelijk ook geholpen wordt en een operatie ondergaat.

Jaklien: Ja.

Sabrina: En wat ik dan met dit framework, het algoritme, dit heet dan de Episode Discovery, die heeft dan deze patroon eruit gehaald die zei, dit komt heel vaak voor.

Jaklien: Ja, maar ja, Ik denk dat je daar In de praktijk niet zo niet zo heel veel mee kan, want ja, dat is logisch. Ze komen dan weer terug. Ze hebben nog steeds bloedverlies en dan komen zij terug. Dan krijgen ze gewoon weer opnieuw een consult met een onderzoek. Dus het is niet zozeer dat door de kliniek en door de ultrasound er weer bloed verlies ontstaat, h? Wat die pijl suggereert, maar het is meer van ja, het bloedverlies houdt aan en dus gaan ze weer opnieuw een afspraak maken en komen ze weer opnieuw bij ons.

Sabrina: Ja. Nou die die pijl dat suggereert eigenlijk van wat de volgende stap is, wat we dus zien in de dataset is dus, deze stap. Dus, dat geeft meer een beetje soort volgorde aan.

Jaklien: Ja, maar precies pijl geeft meestal een volgorde aan, ook binnen de biologie is dat. Maar ja, dat.

Sabrina: Ja.

Jaklien: Dan heeft dit, denk ik, meer te maken met de registratie dan dat dit nou te maken heeft met een logische volgorde van hoe die patint komt . Want die patint heeft gewoon weer opnieuw bloedverlies en die gaat aan de bel trekken. Die zegt, ik kom weer naar jullie kliniek, want het gaat niet goed en op dat moment wordt er eerst een consult aan gemaakt

en dan wordt er een ultra sound neergezet en dan wordt er opnieuw in het dossier geschreven. Heeft nog steeds abnormaal bloedverlies, dus dat is meer een registratie, iets h? Dan dat het lijkt alsof dat op een volgt.

Jaklien: Dan dat dat nou een patroon is waar je, denk ik, iets mee kan.

Sabrina: Oke, ik schrijf wel bij dat het dus meer ligt aan de registratie van die volgorde dan dat dit zozeer het geval is eigenlijk.

Jaklien: Ja, ik denk dat je bij heel veel van dit soort dingen die je ziet, moet je denk ik gewoon heel logisch proberen na te denken hoe je dat zelf gaat bij jou in het ziekenhuis. Ja of bij de huisarts. Je hebt dan een klacht en dan ga je daar naartoe en dan ga je bijvoorbeeld naar de huisarts en dan zeg je: ja, ik heb nog abnormaal bloedverlies. Zij gaan een onderzoek doen, dat onderzoek krijgt gewoon een vaste zorgactiviteit code, dus dat komt meteen in het systeem en na dat onderzoek schrijft de arts je klacht op, maar dan is het niet zo dat die klacht volgt erna. Dat kan natuurlijk wel dat de klacht dan is, op basis van het onderzoek , maar in ons geval is het andersom dat die vrouw met specifiek die klacht komt en daardoor dit onderzoek opnieuw wordt gedaan.

Jaklien: Snap je wat ik bedoel? En dan wordt het gewoon weer opnieuw geregistreerd.

Sabrina: Ja, dat is eigenlijk als die klacht steeds terugkomt, dan wordt er steeds geregistreerd die abnormale bloedverlies.

Jaklien: Precies dan schrijft gewoon arts die zegt van nou, die patint belt , die zegt ja, ik heb nog steeds last van bloedverlies. Ik wil graag weer komen en dan krijgt diegene weer een afspraak. Dan wordt er weer onderzoek gedaan en dan wordt het weer in het dossier opgeschreven. Bloedverlies is er nog steeds. Of het komt weer terug, het is aangehouden, is niet weggegaan of weer opnieuw verschenen.

Sabrina: Ja, dat zou heel goed kunnen, want ik merkte ook in sommige caseIDs van patinten dat sommige activiteiten zich herhalen. Dat er een soort loop in zat en dat is ook wat deze framework soms kunnen vinden. Dus zou kunnen dat dit een soort loop is. Maar, zou je dan zeggen dat dit een patroon is die niet belangrijk is voor de uitkomst?

Jaklien: Nou, ik denk het niet. Kijk, dat laatste blokje natuurlijk wel, want diegene heeft nog abnormaal bloedverlies, dus die krijgt weer een re-interventie. Maar de volgorde hoe dat daar staat, dat is niet van belang. Maar wel gewoon dat iemand weer terugkomt met bloedverlies.

Sabrina: Ja ok, die laatste activiteit is, die klopt dan wel.

Jaklien: Ja.

Sabrina: Oke, zal ik het dan op nee zetten?

Jaklien: Ja, maar niet per se, dus niet die volgorde die daarin staat.

Sabrina: Dus dit is eigenlijk meer een registratie fout.

Jaklien: Het is geen registratie fout, maar het is gewoon de manier hoe registratie gaat. Je doet dan eerst onderzoek en daarna gaat de arts achter de computer zitten en dan gaat hij dit intypen. Maar de activiteit die zitten dan al eerder in, want iemand krijgt een afspraak en bij die afspraak wordt al gezegd, van ok, dat krijgt de patient een afspraak en we gaan een nieuwe ultrasound doen. Nou, dat wordt dan gewoon bij maken van afspraak in het dossier gezet en dan wordt het onderzoek daar gedaan. En dan gaat de arts na dat onderzoek aan de computer zitten. En die schrijft dit dan op.

Sabrina: Ja, oke dat verklaart dan deze volgorde. Dat is wel fijn om het erover te hebben.

Jaklien: Ja, dat.

Sabrina: Nou ja en dan how confident are you for the question above?

Jaklien: Ja, daar ben ik wel confident in.

Sabrina: Ja, dat idee had ik ook dus rond de 9, 8 of?

Jaklien: Ja de 9, 9 is prima.

Sabrina: En waarom?

Jaklien: Ja precies h? Dat registratie van abnormaal bloedverlies vindt plaats na het onderzoek, maar is wel al de aanleiding geweest voor de afspraak en het onderzoek, h? Met die klacht heeft die patint al gebeld.

Sabrina: De registratie van abnormaal bloedverlies vindt al plaats na het onderzoek.

Jaklien: Maar de klacht was al aanwezig voor de afspraak. Want anders, ga je dat onderzoek niet doen?

Sabrina: Ja, waarom heb je anders een afspraak bij de dokter?

Jaklien: Ja.

Sabrina: Oke, ja, ik heb ook met al die patronen een chi-square test gedaan en daaruit bleek ook dat deze niet significant was, dus het kan ook wat je ook zegt. Had je dat ook verwacht, dat die niet significant was? Dus dat gaat nu meer om het gehele patroon en die bleek niet significant te zijn. Dus het heeft geen significante waarde voor de uitkomst, dus of er interventie of geen interventie plaatsvindt.

Jaklien: Nee, dat patroon niet. Maar, als de vraag is in de volgorde hoe het er net stond dan inderdaad niet, maar de klacht zelf wel? Ja, dat weet ik niet. Ik weet niet, dan moet je misschien even noteren dan, want je hebt hier geen ruimte om dat op te schrijven.

Sabrina: Ja ik snap wel wat je bedoelt. Het gaat echt om de volgorde, die klopt hier eigenlijk niet. Het is wel grappig dat je dat zegt, want ik had ook een andere patroon. Die had al die activiteiten dus die 3 dingen . Maar de volgorde was dan weer wel anders en daar zat die abnormaal bloedverlies vooraan, dus die was als eerste. Dus dat is grappig en dat

schiet me nu te binnen, dus ik vind het grappig dat je dat zegt dat. Ja ik begin dan de linkjes te leggen. Maar ja, hoe zou je deze vraag beantwoorden met een ja of een nee?

Jaklien: Ja, als het echt om de volgorde gaat, dan zou ik een nee zeggen, inderdaad. Ja wacht even, even kijken hoor en dan in relatie tot de uitkomst daarbij. Daar gaat het steeds om. Oke, nee, dan zou ik een nee zijn.

Sabrina: Oke, nou ja, waarom eigenlijk?

Jaklien: Ja, dat is eigenlijk hetzelfde antwoord, inderdaad. Misschien kan je wel zeggen wel te verwachten dat het met de klachten een relatie heeft, maar niet zozeer het patroon wat je daar ziet, h?

Sabrina: Dus het abnormaal bloedverlies wel, maar niet het patroon.

Jaklien: Ja.

Sabrina: Oke super.

B.2 Second Process Pattern of Episode Discovery

Sabrina: Zou je verwachten dat dit een belangrijk patroon is voor de uitkomst voor interventie?

Jaklien: Even kijken, ik moet heel even kijken. Wil jij je muis een beetje opzij schuiven? Wat is dat drop out standaard, wat betekent dat? Ik heb geen idee wat dat betekent. Dat is gewoon de cyclus stoornissen. Dat is gewoon de klachten waarmee iemand komt.

Sabrina: Ja, volgens mij wel.

Jaklien: Dan intensieve invasieve therapie. No operating Group. No operating groep 3 no operating groep twee. Wat betekent dan no operating groep 3 en no operating Group twee? Zegt met helemaal niks dit.

Sabrina: Zullen we anders uitgaan van die cyclus, dus dat het de klacht is, want ik wil het wel graag afmaken in verband met de tijd als dat kan?

Jaklien: Ja, dus een patint komt binnen en die krijgt gewoon een DBC code, dat wil zeggen, een diagnose behandel code. Dat is dan cyclus stoornissen en abnormaal bloedverlies valt onder cyclus stoornissen, dus eigenlijk is dat een beetje hetzelfde als dat andere. Ja, iemand komt waarschijnlijk dan weer met een cyclus stoornis die dan nog steeds is gebleven. Je krijgt een ultra sound. Ja en dan wordt het weer genoteerd. Van ja, mevrouw heeft nog steeds abnormaal bloedverlies na de operatie. Als we nu even van cyclus disorder uitgaan, is dat hetzelfde. Maar goed , er staat ook therapie, dat is iets heel anders natuurlijk. Ja, ik denk dat dat dan de novasure is, of niet.

Sabrina: Nee, want dat is nog vr de operatie. Daarvoor staat die before.

Jaklien: Oh ja, die is een beetje verwarrend. Misschien is no operating Group wil dat dan zeggen, h, ik zit maar even hardop te denken. Misschien wordt dan hiermee bedoeld dat de novasure echt de allereerste operatie is dat iemand nog niks anders heeft gehad daarvoor.

Sabrina: Oh, dat zou kunnen. Maar als je kijkt naar het hele patroon zelf, is dat iets wat je zou verwachten dat dit belangrijk is voor de uitkomst ? Om even terug bij de vraag te komen.

Jaklien: Ja, alleen dan die cyclus. Nou ja, ik denk het niet.

Sabrina: En hoe zeker ben je daarover eigenlijk?

Jaklien: Ja, ik vind dit wel iets onzeker omdat er staan ook verschillende dingen bij. Het is een beetje moeilijk omdat daar is alles in n ding is gepropt.

Sabrina: Ja.

Jaklien: Terwijl het misschien hele andere dingen zijn, snap je? Waarop zeg ik dan ja of nee op die cyclus? Want al die verschillende dingen moet je die misschien niet uit elkaar trekken? Want misschien is op het ene antwoord, een ja en op de andere een nee.

Sabrina: De focus is op het hele patroon zelf. Dus met die patroon bedoel ik al die 3 activiteiten bij elkaar.

Jaklien: Ja, dat moet dus allemaal zijn en dan is het een patroon. Ok ja, nou weet je, ik moet effe hier over na gaan denken, want ik kan daar dus nu eigenlijk geen goed antwoord op geven. Dus ja, je kan hier ook neerzetten waar confident 5 of 4 want ik weet gewoon niet zo goed wat hiermee bedoeld wordt.

Sabrina: Nee, dat snap ik, maar dan gaat het meer om het feit dat hier van alles is ingepropt. En dan zet ik dat er ook bij.

Jaklien: Ja zet dan eens maar even bij van ja. Het is nog wat onduidelijk wat al die verschillende factoren inhouden en omdat het een diversiteit is aan factoren.

Sabrina: Nou, Ik heb dan ook hier weer de chi-square test op gedaan. En die blijkt toch wel significant te zijn.

Jaklien: Ja.

Sabrina: Dat patroon, is dat iets wat je zou verwachten? Ja of nee?

Jaklien: Ja, dat kan ik eigenlijk nou niet antwoorden, omdat ik het niet zeker weet, maar dat kan je niet aanklikken, maar goed.

Sabrina: Als we nou uitgaan van die cyclus disorder en dan de rest van de activiteiten. Zou je verwachten dat hij significant is volgens de chi-square test?

Jaklien: Nee. En dan eigenlijk weer een beetje hetzelfde, als dat de eerste

antwoord.

B.3 Third Process Pattern of Episode Discovery

Sabrina: Dit is wat ik al eerder zei, dus nu is het een andere volgorde. Maar nu heb je dus dat de ultrasound van de ingewanden eerst is en daarna komt dan het bezoek. En dan abnormaal bloedverlies. Zou dit een patroon zijn wat je zelf had verwacht en dat belangrijk zou zijn voor de uitkomst?

Jaklien: Maar dit gaat over een interventie toch?

Sabrina: Ja, maar meer van of je verwacht dat dit belangrijk is voor een mogelijke interventie, of niet?

Jaklien: Ja, maar dat is iets anders als er een re-interventie is, h. Dus dan moet ik even heel scherp hebben wat jij nou hier bedoelt?

Sabrina: Ja, dus als je een novasure operatie hebt en dan heb je daarna nog een interventie. Als nou blijkt dat het bloedverlies aanhoudt.

Jaklien: Ja, maar dat noem je een re-interventie. De interventie is de novasure en een re-interventie is wat daarna komt.

Sabrina: Ah, ok, dan heb ik het nu echt over de re-interventie. Ik had begrepen dat novasure nou, dat is gewoon de novasure operatie en daarna heb je nog een mogelijke interventie mocht dat bloedverlies aanhouden. Dat is hoe ik het bedoel.

Jaklien: Maar dat is dus niet in de medische termen. Heb je een interventie , dan is dat eigenlijk de eerste behandeling die je gaat doen. Dus dat is in dit geval de novasure. En als die niet goed is gegaan of het is blijven aanhouden of iemand is niet tevreden, krijgt hij nog een operatie en dat noem je dan een re-interventie.

Sabrina: Ok nou, dan in dit geval heb ik het dan over een mogelijke re-interventie. Dat bedoel ik dan ook met die uitkomst.

Jaklien: Ja ok, dus uitkomst is re-interventie. Dus je krijgt dan voor de novasure een ultrasound, voor de novasure heb je een visite, en voor de novasure heb je abnormaal bloedverlies. Zeg je van, dat bepaalt dus of je een interventie krijgt? Nee, want daarvoor kom je. Dus, dat is nee.

Sabrina: Ok en hoe zeker ben je daarover?

Jaklien: Ja, ben ik heel zeker van, dat heeft niks te maken met of je re-interventie hebt.

Sabrina: Een 9 of een 10 of zelfs een 8?

Jaklien: Ja, doe maar 9.

Sabrina: En waarom?

Jaklien: Ja, die 3 dingen die daar staan zijn de reden voor een novasure. Maar niet de reden voor een re-interventie, h? Als dit zou zijn na de novasure, dan zou het zijn een jaar, maar dit is allemaal voor de novasure.

Sabrina: En dan heb ik weer de chi-square test gedaan en die blijkt dan wel significant te zijn voor een re-interventie. Had je dat verwacht?

Jaklien: Ja, en dat is eigenlijk omdat we weten uit de literatuur dat 10% van de mensen de klacht aanhoudt of weer terugkomt. Dus ja, als iemand met die klacht komt, dan weet ik inderdaad al dat 10% van de mensen nog een keer een operatie moet hebben. Omdat het gewoon niet helpt.

Sabrina: Dat is wel interessant om te horen. Ik had dat ook uit de literatuur gelezen dat 10% een interventie nodig had.

Jaklien: Ja, dus dan verwacht je op zich wel, dat het niet voor iedereen helpt.

B.4 First Process Pattern of Local Process Model

Sabrina: Ja precies nou, nu gaan we door naar de volgende methode om die patronen te achterhalen. Dat is een local process model en dit is de eerste patroon. Het begint met een start. Dit is dan de eerste activiteit. En dan ja, het werkt met tokens, maar dat laat ik even achterwegen. En dan komt de tweede activiteit die daarop volgt. Dat is dan deze. Is dit een patroon waarvan je verwacht dat dit belangrijk is voor een re-interventie, of niet?

Jaklien: Zie krijgt. En een visite. Dan consult. Nee, maar wat ik niet helemaal begrijp, h? Want dit is dus allemaal voor de novasure?

Sabrina: Ja, want er staat niet before.

Jaklien: Dat is natuurlijk logisch, want je weet dat 10% van de mensen het niet helpt en dat ze dan terugkomen.

Sabrina: Ja.

Jaklien: Dus in die zin is het logisch. Ja met dat in het achterhoofd dat je weet dat mensen met een klacht komen, dat 10% altijd terugkomt. Dan verwacht ik wel dat het een ja is. Want ze komen voor die klacht, maar ik vind het is niet zozeer dat die visite en die consult nou bijdragen aan een re-interventie, maar gewoon het feit dat iemand met die klacht komt en dan heb je automatisch een consult en een visite.

Sabrina: Maar dit is nog vr de operatie, dus voor de interventie.

Jaklien: Ja, dat is ook een beetje hetzelfde als wat ik net zei. Iemand heeft een klacht en ik weet nu al op voorhand, voordat iemand berhaupt is gekomen, dat 10% een re-interventie daarvan gaat krijgen. Dus ja, als ik weet dat iemand komt in onze kliniek, dus een visite of een consult

heeft en een telefoongesprek met die klacht. Ja dan weet ik al op voorhand dat 10% inderdaad weer opnieuw een re-interventie gaat krijgen.

Sabrina: Maar je weet niet welke 10%, dus welke patint dat zal zijn?

Jaklien: Nee, maar dat kun je met dit ook niet achterhalen, want iedereen krijgt een visite en een gesprek.

Sabrina: Ja, dat is ook waar. Hoe zou je deze vraag beantwoorden met een ja of een nee?

Jaklien: Tja, vraagstellingen zijn wel erg lastig, moet ik zeggen.

Sabrina: Ja vandaar een interview ook.

Jaklien: Het is een beetje een cirkel, redenaties steeds. Ik verwacht niet dat zij belangrijk zijn, maar het is wel logisch. Dat is eigenlijk het antwoord. Ja, het is niet dat dit nou reden kan zijn waarom ze een interventie hebben, maar het is wel logisch omdat je weet dat 10% van de mensen de klacht weer opnieuw krijgen. Snap je, dus dat moet je eigenlijk zo in het antwoord zetten.

Sabrina: Ok. Ik schrijf dat er zo onder. Hoe zeker ben je dan daarover? Ik heb hem nu op nee gezet, vind je dat ook?

Jaklien: Ja.

Sabrina: OK.

Jaklien: Ja, ik zou het op een 8 zetten.

Sabrina: En dan waarom? Nou eigenlijk wat je net zei, het is een beetje een soort cirkel redenatie en je verwacht al dat 10% terugkomt.

Jaklien: Ja, maar dat patroon zelf draagt niet bij aan een re-interventie.

Sabrina: Ja.

Jaklien: Maar met de redenatie dat 10% sowieso niet heeft geholpen, ja, is het logisch dat dit wel als patroon uit jouw vraag komt. Dus niet de oorzaak, maar meer van ja, dat het logisch. Als het weer aanhoudt komen ze gewoon weer terug, dus iedereen zit daarin.

Sabrina: Ja, ik snap wat je bedoelt. Dat is wel goed om te horen, want dan kan je dat meenemen in mijn discussie. Nou ja, het bleek ook uit de chi-square test dat hij niet significant is. Had je dat verwacht?

Jaklien: Ja.

Sabrina: Nou en dan weer dezelfde. Waarom?

Jaklien: Ja, beetje hetzelfde ja.

B.5 Second Process Pattern of Local Process Model

Sabrina: Dan gaan we naar het tweede patroon toe en dat is een beetje een gekke. Ik ga het uitleggen hoor en dan beginnen we bij start. Er is een gynaecologisch ultrasound en dan gaan we hier naartoe. Controle patint bij de gynaecoloog. En dan gaan we hier naartoe. En wat hij dan kan doen is. Je hebt twee opties: Hij gaat hier naartoe en dan is het klaar. Of hij gaat hier naartoe en dan zit hij in een soort loop met weer controle patint gynaecoloog. Dus hier is een van die loops waar ik het eerder over had.

Jaklien: Ja, dit kan ik me wel voorstellen dat het inderdaad belangrijk is, want iemand komt terug omdat hij niet tevreden is en die komt dan misschien een hele tijd een paar keer terug, want die is gewoon niet voldoende geholpen. Die heeft opnieuw bloedverlies of aanhoudend bloedverlies of gewoon pijn. En dan komt iemand meerdere keren terug en dan kan uiteindelijk besloten worden dat de patint, samen met de arts, nog een operatie gaat doen, dus we doen een re-interventie. Dus die hier , kan ik me wel in vinden.

Sabrina: En, dit is dan weer before die bij de activiteit staat, dus dit is nog vr de interventie.

Jaklien: Je hebt soms gewoon patinten die gewoon een minder goede copingstrategie hebben. Dus die kunnen minder goed met klachten omgaan, die hebben gewoon meer gedoe en meer moeite met hun klacht dan een ander en een ander die kan misschien op een pijnschaal van 8 of op een pijnschaal van 10 zitten. Die andere patint geeft bijvoorbeeld een twee, terwijl ze misschien objectief gezien evenveel bloedverlies hebben. Maar de n zit er gewoon meer mee dan een ander. Dus het is heel erg persoonlijk en individu gevoelig. Hoe ga je met je klacht om in het leven? Ben je negatief ingesteld? Positief ingesteld, dan kan je er goed mee dealen. Heb je misschien thuis hele goede begeleiding of opvang? Dus er zijn veel meer randvoorwaarden die te maken hebben met een re-interventie.

Sabrina: Ja.

Jaklien: Het individu, het type mens en hoe een copingstrategie en omgeving . Dingen die maken of iemand wel of niet een re-interventie heeft. Dat zou dan mijn verklaring zijn voor deze loop die je hebt. Dus ik kan me voorstellen dat iemand die weinig steun heeft vanuit thuis of die heel negatief is of h die niet zo goed met pijn of klachten kan omgaan dat die sneller terugkomt en dus ook na de novasure nog steeds ontevreden blijft. Ja snap je wat ik bedoel of niet?

Sabrina: Ja, ik snap wat je bedoelt, want dat zijn eigenlijk omgevingsfactoren die ook uitmaken.

Jaklien: Ja omgevingsfactoren en dus de intrinsieke factoren van de patint. Die, denk ik, ervoor kunnen zorgen dat en dat is wel interessant. Ja, want het is wel bekend dat mensen een slechte copingstrategie dus niet goed om kunnen gaan met een klacht en komen dus veel vaker terug en hebben dus ook meer ingrepen. Omdat ze gewoon niet zo snel tevreden zijn .

Sabrina: Ja op zich klinkt dat heel logisch ook.

Jaklien: Ja, dus het is wel interessant.

Sabrina: Hoe confident ben je over die patroon?

Jaklien: Ja, ik denk, dat ik een 8 kan zeggen.

Sabrina: Nou, ik heb natuurlijk weer de chi-square test gedaan en hij bleek ook significant te zijn, dus dit is heel interessant om te horen.

Jaklien: Dat is wel echt leuk. Hoor dat je dat zo eruit haalt.

Sabrina: Ja ik heb hem ook op een ja gezet. En waarom?

Jaklien: Vanwege de omgeving van de patint kan je verwachten dat er een interventie of niet plaats vindt. Dus dat die patint steeds terugkomt, h ? Dus steeds weer opnieuw naar die kliniek op zoek. En dat dat dus iets zegt over hoe iemand met zijn klacht omgaat. Waarschijnlijk kan diegene dan toch niet zo goed met zijn klacht omgaan of heeft hij niet zo'n ondersteunende omgeving? En dan verwacht je misschien ook sneller dat hij een re-interventie krijgt.

B.6 Third Process Pattern of Local Process Model

Sabrina: Ok, dan gaan we naar het derde patroon van dit algoritme. En hier zie je eigenlijk iets leuks. Tenminste, dat vind ik zelf leuk. Nou ja, beginnen weer bij start komen we hier uit en dan zie je een soort parallel van activiteiten en het is weer helaas die drop out. Dus laten we even uitgaan van cyclus.

Jaklien: OK ja. Waarom staan die twee losse blokken, zal ik maar zeggen, is de een positief en de andere negatief of wat is het?

Sabrina: Je bedoelt deze twee blokken?

Jaklien: Ja, die onder elkaar staan?

Sabrina: Nou ja, we noemen dit parallellisme, dus ze gebeuren tegelijk of vlak na elkaar.

Jaklien: Ok, op die manier.

Sabrina: Vandaar dat ze zo op deze manier zijn afgebeeld. En is dat een patroon waarvan je denkt van: dat is belangrijk voor een mogelijke interventie voor de uitkomst?

Jaklien: Dat kan ik niet inschatten, dat moet ik echt vragen aan de clinicus, omdat ik geen idee heb.

Sabrina: OK.

Jaklien: Ja, ik heb geen idee. Ik weet ook niet wat... Ik denk, nou,

therapeutic light. Ik denk, wat is dat? Ik heb geen idee wat dat is. Ja, het zegt mij gewoon echt helemaal niks, dit.

Sabrina: Zal ik het anders op een 4 zetten?

Jaklien: Ja, dat is echt maar even zo. Ik weet het niet. Mijn advies zou zijn dat je even kijkt van wat het nou precies is. Dit is echt iets wat de gynaecoloog moet aangeven, omdat het echt klinisch gerelateerde zaken zijn waarvan ik dat dan echt niet durf te zeggen.

Sabrina: Nee, dat snap ik, maar dan zet ik hem even zo neer, maar hij was ook niet significant.

Jaklien: Ok.

Sabrina: Ja, zou je dat verwachten? Dat is ook een moeilijke vraag nu denk ik.

Jaklien: Nee, want wat is dat light? Ik weet niet wat dat betekent?

Sabrina: Nou ja, ik hoopte eigenlijk dat jij dat zou kunnen uitleggen, want ik heb ook de dataset zo gekregen.

Jaklien: Grappig wat dit betekent.

Sabrina: Ja, daar stond dit wel in. Ik moet er wel bij zeggen dat het maar 13 keer voorkwam, dus het is wel een heel uitzonderlijk activiteit.

Jaklien: Ja, kan me voorstellen. No clinical recording no light.

Sabrina: Het is wel een beetje een gekke, moet ik ook eerlijk toegeven.

Jaklien: Dit is dus after, dus dit wil dan weer zeggen na de novasure.

Sabrina: Ja, inderdaad na de novasure, goed goed gezien!

Jaklien: Dit is dus niet tijdens h. Dit is echt na, want je hebt natuurlijk ook nog peri-operatieve dingen zoals de power en de tijdsduur van de novasure, want had ik nog ook nog wel verwacht dat dat misschien ergens uit zou komen? Maar dit is echt na de novasure.

Jaklien: Ja, maar ik denk na de novasure. Wat is dat dan, de operating light en therapeutic light? Dus ik vind het zo vreemd, want na de operatie is het gewoon klaar. Na de novasure dus. Nou goed, misschien moeten we even allebei kijken naar die dataset. Ik weet niet of ze het heeft uitgelegd hoe dat werkt, h? Dat is een speciaal programma, dat ziet de que en daarmee trek je dus uit het EPD klinische gegevens. Dat is aan de hand van een query die ik dan schrijf, maar hij neemt ook automatisch dingen mee die daaraan gekoppeld zijn. Dus als ik bijvoorbeeld vraag novasure, dan neemt die de ingreep novasure. Dan neemt hij bijvoorbeeld automatisch ook mee: exacte tijdstip en misschien nog andere details van de operatie die ik niet specifiek heb gevraagd. Ik weet dat ik toen tegen Xixi heb gezegd: van ja, zal ik dat allemaal weghalen en alleen maar novasure doen? Dat laten we alleen waarvan ik denk dat het relevant is, maar Xixi zei toen van, nee, laat maar gewoon alles erin.

Sabrina: Ja, ik weet wel dat dit alleen van een kolom is, want ik had dat ook inderdaad gezien en dat was het kolom met activiteiten. Ik moet anders weer erbij pakken, maar vanwege de tijd zullen we doorgaan naar de volgende. Ja, en waarom?

Jaklien: Ja, dit kan ik eigenlijk niet beantwoorden.

Sabrina: Ja, ik zet het er ook bij. Vanwege de onduidelijke termen.

Jaklien: Ja.

B.7 First Process Pattern of IMPresseD Framework

Sabrina: En nu gaan we naar het volgende framework, en dat is de impressed framework. En ja, je ziet die gekleurde bolletjes en hier staat dan de legenda. Dit is dan ook weer een soort parallellisme of eigenlijk vlak na elkaar gebeurt en dan dus ook weer before.

Jaklien: Dat is eigenlijk weer precies als helemaal in het begin. Gezien de 10% die interventie krijgt is het logisch, maar dit draagt niet bij aan de interventie dus.

Sabrina: Dus het is meer een nee hoor ik.

Jaklien: Ja, eigenlijk wel, ja. Als we kijken naar de oorzakelijke verbanden dan zou ik hier een nee zeggen. Logisch gezien, zou ik ja zeggen, omdat je weet dat 10% gewoon weer opnieuw terugkomt omdat het gewoon niet helpt. Dus ja, misschien moet je dat er ook bij zetten, dus oorzakelijk verwachting is nee, maar logisch, ja, omdat 10% terugkomt.

Sabrina: En hoe zeker ben je erover?

Jaklien: Ja, dan ben ik wel zeker over, ja.

Sabrina: 8 of 9?

Jaklien: Ja, een 8.

Sabrina: Nou, ik heb dan weer de chi-square test gedaan die zei dat het wel significant is.

Jaklien: Ja, daar sta ik wel achter. Dus wel logisch ook vanwege die 10%, maar het is niet oorzakelijk naar mijn idee.

Sabrina: Ja, want het komt wel vaak voor bij die mensen van die 10% die dan terugkomt.

Jaklien: Ja, die komen met een klacht en je weet al op voorhand dat het bij 10% niet helpt, maar dat heeft niks te maken met die klacht en dat er een onderzoek is. Maar gewoon omdat ze met die klacht komen. Ik weet dat het bij 10% niet helpt.

B.8 Second Process Pattern of IMPresseD Framework

Sabrina: Dan gaan we door naar de volgende. Hier is weer een after, dus na de interventie, na de novasure operatie. Ik zie een nursing day en daarna kunnen ze beide voorkomen, die grijze bolletjes, of alleen de een gebeurt en de ander niet. Dus wat er daarna kan gebeuren, is consult via de telefoon en nog een teleconsult, of een teleconsult of een consult?

Jaklien: Ja, dus dit is dan na de novasure.

Sabrina: Ja.

Jaklien: Ja ok, iedereen krijgt gewoon standaard na de novasure een telefoontje of een consult. Ik denk nee, maar goed, als tijdens dat gesprek wordt gezegd van ja, ik heb nog steeds heel veel klachten, ja dan kan het zo zijn dat iemand dus nog een keer terugkomt omdat hij dan een re-interventie wil. Dus iedereen krijgt dit, dus in dat opzicht is het niet belangrijk voor de uitkomst. Maar als daar tijdens dat gesprek wordt gezegd, ja, ik ben ontevreden, ja, dan wel. Snap je wat ik bedoel of niet?

Sabrina: Ja, ik snap wat je bedoelt, maar het hele patroon, dat is niet wat je verwacht?

Jaklien: Ja, iedereen krijgt gewoon standaard zo'n gesprek. Of je nou een interventie krijgt ja of nee. Maar wat er tijdens het gesprek wordt gezegd, ja, dat is wel belangrijk, maar niet zozeer dat het gesprek er is, want dat krijgt gewoon iedereen.

Sabrina: Ja, dus als ik kijk naar het hele patroon, dan komt het eigenlijk wel bij iedereen voor om dat iedereen dat krijgt.

Jaklien: Nou ja, die stelling zijn dus niet zo zwart-wit.

Sabrina: Nee, ik snap wat je bedoelt. Wil je hem dan nu op een nee zetten van dat het niet belangrijk is voor een interventie?

Jaklien: Ik zet maar even op. Ja, en dan is het een 5, 6.

Sabrina: Ja, en dan schrijf ik erbij dat iedereen zon gesprek krijgt.

Jaklien: Iedereen krijgt standaard een consult na de novasure. Dus in dat opzicht is de consult niet voorspellend voor een re-interventie. Maar wat er gezegd wordt tijdens het gesprek, dus de inhoud van het gesprek, kan wel bepalend zijn voor een re-interventie, want als iemand dan zegt: ik heb nog steeds klachten. Dus ja, misschien moet je confidence 6 of 7 van maken.

Sabrina: Het was wel significant.

Jaklien: Ja oke, nou misschien is dat dan door dat iemand zegt: ja, ik ben helemaal niet tevreden of ik zie nu al dat het niet werkt. Dat zou kunnen.

Sabrina: Dus je had dat niet verwacht en omdat iemand zegt dat hij niet tevreden is?

Jaklien: Nou ja, de inhoud van het telefoongesprek verwacht ik wel dat het van invloed is. Maar het telefoongesprek an sich niet, omdat iedereen een telefoongesprek krijgt.

Sabrina: Ja, maar we kijken nu echt puur naar de activiteiten, dus niet de inhoud van de gesprekken.

Jaklien: Niet relevant dan, want ja, weet je dat is, dat is gewoon standaardzorg. Dus het draagt niet bij aan een re-interventie.

Sabrina: Nee, precies omdat het een standaardprocedure is voor elke patint, oke?

Jaklien: Ja, het is gewoon een controle eigenlijk, he?

B.9 Third Process Pattern of IMPresseD Framework

Sabrina: Dan gaan we door naar de volgende. Hier is weer een after, dus na de interventie, na de Novasure-operatie. Ik zie een nursing day en daarna kunnen ze beide voorkomen, die grijze bolletjes, of alleen een gebeurt en de ander niet. Dus wat er daarna kan gebeuren, is consult via de telefoon en nog een teleconsult of een teleconsult of een consult?

Jaklien: Ja, dus dit is dan na de Novasure.

Sabrina: Ja.

Jaklien: Ja ok, iedereen krijgt gewoon standaard na de Novasure een telefoontje of een consult. Ik denk niet dat het een bepalende factor is voor een re-interventie. Maar goed, als tijdens dat gesprek wordt gezegd van ja, ik heb nog steeds heel veel klachten, dan kan het zo zijn dat iemand dus nog een keer terugkomt omdat hij dan een re-interventie wil. Dus iedereen krijgt dit, dus in dat opzicht is het niet belangrijk voor de uitkomst. Maar als daar tijdens dat gesprek wordt gezegd, ja, ik ben ontevreden, dan wel. Snap je wat ik bedoel of niet?

Sabrina: Ja, ik snap wat je bedoelt, maar het hele patroon dat is niet wat je verwacht?

Jaklien: Ja, iedereen krijgt gewoon standaard zo'n gesprek. Of je nou een interventie krijgt of niet. Maar wat er tijdens het gesprek wordt gezegd , dat is wel belangrijk, maar niet zozeer dat het gesprek er is, want dat krijgt gewoon iedereen.

Sabrina: Ja, dus als ik kijk naar het hele patroon, dan komt het eigenlijk wel bij iedereen voor dat iedereen dat krijgt.

Jaklien: Nou ja, die stelling is dus niet zo zwart-wit.

Sabrina: Nee, ik snap wat je bedoelt. Wil je hem dan nu op een nee zetten, dat het niet belangrijk is voor een interventie?

Jaklien: Ik zet maar even op. Ja, en dan is het een 5 of 6.

Sabrina: Ja, en dan schrijf ik erbij dat iedereen zon gesprek krijgt.

Jaklien: Iedereen krijgt standaard een consult na de Novasure. Dus in dat opzicht is het consult niet voorspellend voor een re-interventie. Maar wat er gezegd wordt tijdens het gesprek, dus de inhoud van het gesprek, kan wel bepalend zijn voor een re-interventie. Want als iemand dan zegt: ik heb nog steeds klachten, dan kan dat wel invloed hebben. Dus ja, misschien moet je confidence 6 of 7 van maken.

Sabrina: Het was wel significant.

Jaklien: Ja ok, nou misschien is dat dan door dat iemand zegt, ja, ik ben helemaal niet tevreden of ik zie nu al dat het niet werkt. Dat zou kunnen.

Sabrina: Dus je had dat niet verwacht en omdat iemand zegt dat hij niet tevreden is?

Jaklien: Nou ja, de inhoud van het telefoongesprek verwacht ik wel dat het van invloed is. Maar het telefoongesprek an sich niet, omdat iedereen een telefoongesprek krijgt.

Sabrina: Ja, maar we kijken nu echt puur naar de activiteiten, dus niet de inhoud van de gesprekken.

Jaklien: Niet relevant dan, want ja, weet je, dat is gewoon standaardzorg. Dus het draagt niet bij aan een re-interventie.

Sabrina: Nee, precies omdat het een standaardprocedure is voor elke patint. OK?

Jaklien: Ja, het is gewoon een controle eigenlijk, h?

Sabrina: Ja, dat vond ik ook wel een gekke. Dit is de laatste patroon hoor, en dan komen we bij de laatste vraag. Nou, weer een soort parallellisme , dus dat het vlak na elkaar gebeurt of misschien zelfs tegelijk. En dan heb je controle. Bovendien is het weer een after, dus na de interventie . Dus controle, patint bij gynaecoloog, repeat polikliniek visit, en dan repetition. Dit vond ik wel een beetje gek, want het is alleen het woord repetition.

Jaklien: Oke, als iemand wat vaker komt of zo.

Sabrina: Dus wat zou je verwachten?

Jaklien: Ja, eigenlijk wel, dus eigenlijk een beetje hetzelfde antwoord als met dat loopje. Zou ik weer een 8 zetten en dan zou ik zeggen, hetzelfde antwoord als dat met dat loopje wat je eerder liet zien.

Sabrina: Ok. Hij bleek ook significant te zijn voor de uitkomst. Had je dat verwacht?

Jaklien: Ja.

Sabrina: En waarom?

Jaklien: Ja, nou, omdat we wel die ervaring hebben in de kliniek, h? Dus in de kliniek heb je de ervaring wel dat patinten een bepaalde karakter hebben of copingstrategien, of misschien wellicht te weinig ondersteuning krijgen thuis. Maar in ieder geval, een copingstrategie is wel bekend dat dat iets zegt over hoe vaak iemand terugkomt. Ik heb ook weer een arts horen zeggen: ik weet al op voorhand aan het TYPE mevrouw dat die gaat terugkomen.

B.10 Representation of Process Pattern Discovery Methods

Sabrina: Dat is ook wel interessant. Zijn we bij de laatste vraag aangekomen en het gaat meer over de representatie, want je hebt drie modellen gezien en elk heeft een eigen manier van het weergeven van die processpatronen. Dit is dan de activiteit die op die dag bij de patint heeft plaatsgevonden. Dus activiteit A is bijvoorbeeld een bezoek aan de polikliniek en op dag twee hebben die twee activiteiten dus tegelijk of na elkaar plaatsgevonden. Patint 3 heeft dan weer activiteit A, C en dan B. Dit is dan hoe model n het zou weergeven, die drie activiteiten voor elke patint. Dit is hoe model twee het zou weergeven, dus eigenlijk kunnen we al die drie activiteiten in een model weergeven. Model drie heeft het dan op deze manier en dan mijn vraag aan jou is: welke representatie heeft jouw voorkeur? Ik ga een klein beetje uitzoomen, zodat je alle drie in beeld hebt.

Jaklien: Ja ja ja, moet ik wel even kijken hoor.

Sabrina: Dus in die hele dataset bij wijze van heb je 35 keer activiteit A. En 20 daarvan staan dan op deze positie in een patroon.

Jaklien: Ok, ja dus dat is meer voor jullie dan, h, want dan zou ik dat zelf niet gebruiken. En dan heb je een black box. En dan heb je B en C. Ja, wat ik denk wat onduidelijker is, h, met die getallen dan met name, en je hebt natuurlijk die box en lege witte bolletjes waarvan ik denk van ja, wat is dat dan? Maar goed.

Sabrina: Ja precies. Dat zijn meer termen en symbolen vanuit de process mining.

Jaklien: Dat is ook prima, maar wel wat rommeliger omdat ik denk van ja, waarom is dat? Aan de andere kant, kan het wel helpen bij ons misschien om te begrijpen, omdat je dan echt ziet van nou: er zijn twee black boxes h? Tenminste, ik neem aan dat dat een black box is of niet?

Sabrina: Nou ja, hoe model twee werkt is met een token-systeem, dus dan

gaat hier een token in en dan heb je hier twee tokens die de black box geven. Twee tokens die komen dan hierin en als die token hier in gaat, dan komt hij hier terecht. Dan gaat hij weer terug in de black box.

Jaklien: Wat wil je met een token? Wat is dat?

Sabrina: Ja, bijna een soort muntje dat door het proces gaat en dan zie je dus als die langs A gaat dan gebeurt activiteit A. Gaat hij langs B, dan gebeurt activiteit B. Maar dat is meer een terminologie van dit model specifiek.

Jaklien: Ja ok.

Sabrina: Vandaar dat ik ook een representatievraag erin wou hebben.

Jaklien: Ja ja ok en dan hebben we model drie.

Sabrina: Ja.

Jaklien: Dat is eigenlijk precies hetzelfde, alleen dan geen bolletjes, maar gewoon gelijk ABC, ja?

Sabrina: Dat representeert eigenlijk alle drie de patinten ook. Want met model drie heb je dat A begint eerst, want dat begint bij elke patint, en dan heb je dat B daarna kan gebeuren of C. En dat kan alleen of dat kan tegelijk.

Jaklien: Maar bij model drie, is het dan wel ietsjes onduidelijker dan? Even kijken. Ok, ja nou ik moet even goed kijken, dus het eerste model is echt de patint. Dan heb je de patint centraal? En bij model drie heb je eigenlijk de activiteit centraal, zo moet ik het zien, h? Want daar zie ik wat een patint is.

Sabrina: Oh nee, dat is voor model n. Model n heeft drie modellen nodig om de drie patinten te weergeven. Patint n die wordt op die manier uitgebeeld. En voor patint twee wordt hij op die manier uitgebeeld en patint drie moet op die manier. Maar dat is meer om aan te geven bij welk model elke patint hoort. Maar dat hoort normaal niet zo. Er wordt dan niet de patint bij geschreven.

Jaklien: Nee, ok, maar dan heb je dus bij het eerste model dus verschillende vormen van presentatie, maar bij model drie heb je die dus niet, want als we even naar model drie kijken dan heeft die dus niet.

Sabrina: Ja.

Jaklien: Dan zou ik zeggen, ok, iemand heeft een activiteit A en die kan dan daarna B krijgen of C? Maar het is mij niet zo logisch dat hij dus ook B/C tegelijk kan hebben. Want ik volg de pijl en dan denk ik, ok, ja hierna krijgt iemand B, maar je kan ook C krijgen. Maar het is mij dan niet duidelijk dat iemand ook bij B/C kan krijgen, dus een model drie vind ik dan wel wat onduidelijker. Ja en bij model n vind ik het ook een beetje onduidelijk dat als iemand twee dingen tegelijk krijgt, dit vind ik ook bij het eerste model en het tweede model eigenlijk allebei er niet goed uitkomen, want bij model n denk ik ook van ja: Ik ben een patint en ik kan A krijgen en dan krijg ik B ofwel C. Dus dan komt

eigenlijk bij alle drie de modellen niet zo heel duidelijk naar boven. En bij patint twee denk ik, ok, krijg eerst A en dan krijg ik B en dan C . En dat klopt!

Sabrina: Ja, dan bedoel je dan wanneer je op dag twee B en C krijgt?

Jaklien: Ja. Je zou bijna zeggen dan alle drie de modellen eigenlijk niet zo duidelijk zijn en dat komt door die pijlen. Wij zijn altijd heel erg gewend vanuit de biologie dat je echt het ene en het andere opvolgt en nu zie ik niet goed dat het ook nog een optie is dat iemand die twee dingen allebei kan krijgen. Maar ja, misschien moet je dan een legende toevoegen dat de een dit betekent en de ander dat.

Sabrina: Ja.

Jaklien: Ja ik lees het als: je krijgt of van het ene of het andere, maar niet allebei.

Sabrina: Maar je hebt dan geen voorkeur voor n van deze drie modellen als je moest kiezen?

Jaklien: Ja, dan zou ik, denk ik, voor model n gaan, omdat bij model twee: hoe moet ik dan patint twee lezen? En model drie die is dus niet apart gesplitst, die is gewoon of voor mij.

Sabrina: Ja dat klopt, model drie is, in die zin wat je ook al eerder zei, wat onduidelijker. Dat noemen we ambigue.

Jaklien: Ja ok.

Sabrina: Ja, daar passen al deze tracen hier in, dus al die drie patinten. Maar dat is inderdaad wat je ook zei, dat maakt het ook onduidelijker.

Jaklien: Ja, maar goed, dan zie je de volgorde niet meer. Want als jij zegt , op dag n krijgt iemand A op dag twee krijgt die B en op dag drie C. Dat is niet logisch met model drie.

Sabrina: Ja, Episode Discovery is wat ambigue. Maar het is wel simpeler daardoor. Het is dus niet heel logisch, maar je moet dan maar net de terminologie kennen. Vandaar dat ik deze vraag ook erin heb zitten om te kijken, welk model heeft jouw voorkeur en dus welk framework?

Jaklien: Ja.

Sabrina: Dan zijn we eigenlijk aan het eind gekomen van het interview en ook van alle vragen.

C. Quantitative Analysis Code

C.1 Code for Decision Tree Model Process Pattern excluded

Below the code for the Decision Tree Model can be found without any process patterns in the dataset. This will be used as a baseline.

```
# Import necessary libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import KFold
from sklearn.metrics import accuracy_score, precision_score, recall_score,
    f1_score, roc_curve, auc


# Disable warnings
import warnings
warnings.filterwarnings('ignore')

#Import CSV file
df = pd.read_csv('../data/Top 10 freq events and caseIDs.csv', sep=';')

#Display first 10 rows
df.head(10)

# Define columns of interest
columns_of_interest = [
    'After_Consult by phone',
    'Before_Ultrasound of the abdominal organs.',
    'Before_First outpatient clinic visit.',
    'After_Repeat-pole clinic visit.',
    'Before_Repeat-pole clinic visit.',
    'Before_Consult by phone',
    'Before_Abnormal blood loss',
    'After_Teleconsult (excluding screen to screen image contact, see
    190019).',
    'Before_First outpatient visit',
    'Before_Repetition'
]

# Create X with the specified columns
X = df[columns_of_interest]
# Display the first few rows of X
print(X.head())
y = df['Intervention']

# Initialize KFold
kf = KFold(n_splits=10, shuffle=True, random_state=1)
```

```
# Lists to store evaluation metrics
accuracies = []
precisions = []
recalls = []
f1_scores = []
auc_scores = []

# Arrays to store ROC curve data
mean_fpr = np.linspace(0, 1, 100)
tprs = []
aucs = []

# Perform k-fold cross-validation
for train_index, test_index in kf.split(X):
    X_train, X_test = X.iloc[train_index], X.iloc[test_index]
    y_train, y_test = y.iloc[train_index], y.iloc[test_index]

    # Initialize and fit the model
    model = DecisionTreeClassifier(class_weight='balanced', random_state=1)
    model.fit(X_train, y_train)

    # Predict probabilities and evaluate
    y_pred_proba = model.predict_proba(X_test)[:, 1]
    y_pred = model.predict(X_test)

    # Calculate evaluation metrics
    accuracies.append(accuracy_score(y_test, y_pred))
    precisions.append(precision_score(y_test, y_pred, average='weighted'))
    recalls.append(recall_score(y_test, y_pred, average='weighted'))
    f1_scores.append(f1_score(y_test, y_pred, average='weighted'))

    # Calculate ROC curve
    fpr, tpr, _ = roc_curve(y_test, y_pred_proba)
    tprs.append(np.interp(mean_fpr, fpr, tpr))  # Interpolate TPR values
    tprs[-1][0] = 0  # Ensure the ROC curve starts at (0,0)
    roc_auc = auc(fpr, tpr)
    aucs.append(roc_auc)

    # Store AUC score for this fold
    auc_scores.append(roc_auc)

# Compute mean and standard deviation for metrics
mean_accuracy = np.mean(accuracies)
std_accuracy = np.std(accuracies)

mean_precision = np.mean(precisions)
std_precision = np.std(precisions)

mean_recall = np.mean(recalls)
std_recall = np.std(recalls)

mean_f1 = np.mean(f1_scores)
std_f1 = np.std(f1_scores)

mean_auc = np.mean(auc_scores)
```

```
std_auc = np.std(auc_scores)

print(f'Cross-Validation Accuracies: {accuracies}')
print(f'Mean Accuracy: {mean_accuracy:.3f}  {std_accuracy:.3f}')
print(f'Mean Precision: {mean_precision:.3f}  {std_precision:.3f}')
print(f'Mean Recall: {mean_recall:.3f}  {std_recall:.3f}')
print(f'Mean F1 Score: {mean_f1:.3f}  {std_f1:.3f}')
print(f'Mean AUC Score: {mean_auc:.3f}  {std_auc:.3f}')

# Plot the results
plt.figure(figsize=(12, 10))

# Accuracy per Fold
plt.subplot(2, 2, 1)
plt.plot(range(1, len(accuracies) + 1), accuracies, marker='o', linestyle='
    -', color='b')
plt.title('Accuracy per Fold')
plt.xlabel('Fold')
plt.ylabel('Accuracy')
plt.ylim(0, 1)

# Precision per Fold
plt.subplot(2, 2, 2)
plt.plot(range(1, len(precisions) + 1), precisions, marker='o', linestyle='
    -', color='r')
plt.title('Precision per Fold')
plt.xlabel('Fold')
plt.ylabel('Precision')
plt.ylim(0, 1)

# Recall per Fold
plt.subplot(2, 2, 3)
plt.plot(range(1, len(recalls) + 1), recalls, marker='o', linestyle='-',
    color='g')
plt.title('Recall per Fold')
plt.xlabel('Fold')
plt.ylabel('Recall')
plt.ylim(0, 1)

# F1 Score per Fold
plt.subplot(2, 2, 4)
plt.plot(range(1, len(f1_scores) + 1), f1_scores, marker='o', linestyle='-'
    , color='purple')
plt.title('F1 Score per Fold')
plt.xlabel('Fold')
plt.ylabel('F1 Score')
plt.ylim(0, 1)

plt.tight_layout()
plt.show()

# Plot mean ROC curve
plt.figure(figsize=(8, 6))
plt.plot([0, 1], [0, 1], linestyle='--', color='r', label='Random Guessing'
    )
mean_tpr = np.mean(tprs, axis=0)
```

```
mean_auc = auc(mean_fpr, mean_tpr)
std_auc = np.std(aucs)
plt.plot(mean_fpr, mean_tpr, color='b', label=f'Mean ROC (AUC = {mean_auc
    :.2f}  {std_auc:.2f})')
plt.fill_between(mean_fpr, mean_tpr - std_auc, mean_tpr + std_auc, color='
    blue', alpha=0.2)
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Mean ROC Curve')
plt.legend(loc='lower right')
plt.show()
```

C.2 Code for Decision Tree Model Process Pattern included

The import of the necessary libraries are the same as for the pattern excluded code in C.1, hence they won't be shown here. Furthermore, the dataset needs to be changed according to the dataset of the process pattern in the import of the CSV file. In addition, the dataset contains an extra column namely, the process pattern column. Therefore, an additional column is added to the features of the model. Lastly, the same code is used as in C.1 for training the model and plotting the graphs, hence it won't be shown here.

```
#Import CSV file and change accordingly to the process pattern dataset
df = pd.read_csv('../data/PROCESS_PATTERN_DATASET.csv', sep=';')

#Display first 10 rows
df.head(10)

# Define columns of interest
columns_of_interest = [
    'After_Consult by phone',
    'Before_Ultrasound of the abdominal organs.',
    'Before_First outpatient clinic visit.',
    'After_Repeat-pole clinic visit.',
    'Before_Repeat-pole clinic visit.',
    'Before_Consult by phone',
    'Before_Abnormal blood loss',
    'After_Teleconsult (excluding screen to screen image contact, see
    190019).',
    'Before_First outpatient visit',
    'Before_Repetition',
    'Pattern_Included'
]

# Create X with the specified columns
X = df[columns_of_interest]
# Display the first few rows of X
print(X.head())
```

```
y = df['Intervention']
```

C.3 Code for Logistic Regression Model Process Pattern excluded

Below the code for the Logistic Regression Model can be found without any process patterns in the dataset. This will be used as a baseline.

```
# Import necessary libraries
import pandas as pd
import numpy as np
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split, GridSearchCV, KFold
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score, precision_score, recall_score,
    f1_score,  roc_auc_score, roc_curve, make_scorer, auc
import matplotlib.pyplot as plt

# Disable warnings
import warnings
warnings.filterwarnings('ignore')

#Import CSV file
df = pd.read_csv('../data/Top 10 freq events and caseIDs.csv', sep=';')

#Display first 10 rows
df.head(10)

# Define columns of interest
columns_of_interest = [
    'After_Consult by phone',
    'Before_Ultrasound of the abdominal organs.',
    'Before_First outpatient clinic visit.',
    'After_Repeat-pole clinic visit.',
    'Before_Repeat-pole clinic visit.',
    'Before_Consult by phone',
    'Before_Abnormal blood loss',
    'After_Teleconsult (excluding screen to screen image contact, see
    190019).',
    'Before_First outpatient visit',
    'Before_Repetition'
]

# Create X with the specified columns
X = df[columns_of_interest]
# Display the first few rows of X
print(X.head())
y = df['Intervention']

# Standardize the features
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
```

```
# Define the logistic regression model
log_reg = LogisticRegression()

# Define the hyperparameters for tuning
param_grid = {
    'penalty': ['l1', 'l2', 'elasticnet', 'none'],
    'C': [0.01, 0.1, 1, 10, 100],
    'solver': ['lbfgs', 'saga', 'liblinear'],
    'max_iter': [100, 200, 300]
}

# Initialize KFold cross-validation
kf = KFold(n_splits=10, shuffle=True, random_state=1)

# Lists to store the evaluation metrics
accuracies = []
precisions = []
recalls = []
f1_scores = []
auc_scores = []
tprs = []
aucs = []
mean_fpr = np.linspace(0, 1, 100)  # Common FPR values for all folds

# Perform k-fold cross-validation
for train_index, test_index in kf.split(X_scaled):
    X_train, X_test = X_scaled[train_index], X_scaled[test_index]
    y_train, y_test = y.iloc[train_index], y.iloc[test_index]

    # Grid search for hyperparameter tuning
    grid_search = GridSearchCV(estimator=log_reg, param_grid=param_grid, cv
    =5, verbose=1, n_jobs=-1)
    grid_search.fit(X_train, y_train)

    # Get the best model
    best_log_reg = grid_search.best_estimator_

    # Predict and evaluate
    y_pred = best_log_reg.predict(X_test)
    y_pred_proba = best_log_reg.predict_proba(X_test)[:, 1]

    accuracies.append(accuracy_score(y_test, y_pred))
    precisions.append(precision_score(y_test, y_pred, average='weighted'))
    recalls.append(recall_score(y_test, y_pred, average='weighted'))
    f1_scores.append(f1_score(y_test, y_pred, average='weighted'))
    auc_scores.append(roc_auc_score(y_test, y_pred_proba))

    # Compute ROC curve and AUC
    fpr, tpr, _ = roc_curve(y_test, y_pred_proba)
    tprs.append(np.interp(mean_fpr, fpr, tpr))  # Interpolate TPR values
    tprs[-1][0] = 0  # Ensure the ROC curve starts at (0,0)
    aucs.append(auc(fpr, tpr))

# Compute mean and standard deviation
mean_accuracy = np.mean(accuracies)
std_accuracy = np.std(accuracies)
```

```
mean_precision = np.mean(precisions)
std_precision = np.std(precisions)

mean_recall = np.mean(recalls)
std_recall = np.std(recalls)

mean_f1 = np.mean(f1_scores)
std_f1 = np.std(f1_scores)

mean_auc = np.mean(auc_scores)
std_auc = np.std(auc_scores)

print(f'Cross-Validation Accuracies: {accuracies}')
print(f'Mean Accuracy: {mean_accuracy:.3f}  {std_accuracy:.3f}')
print(f'Mean Precision: {mean_precision:.3f}  {std_precision:.3f}')
print(f'Mean Recall: {mean_recall:.3f}  {std_recall:.3f}')
print(f'Mean F1 Score: {mean_f1:.3f}  {std_f1:.3f}')
print(f'Mean AUC Score: {mean_auc:.3f}  {std_auc:.3f}')

# Plot the results
plt.figure(figsize=(12, 12))

# Accuracy per Fold
plt.subplot(2, 2, 1)
plt.plot(range(1, len(accuracies) + 1), accuracies, marker='o', linestyle='
    -', color='b')
plt.title('Accuracy per Fold')
plt.xlabel('Fold')
plt.ylabel('Accuracy')
plt.ylim(0, 1)

# Precision per Fold
plt.subplot(2, 2, 2)
plt.plot(range(1, len(precisions) + 1), precisions, marker='o', linestyle='
    -', color='r')
plt.title('Precision per Fold')
plt.xlabel('Fold')
plt.ylabel('Precision')
plt.ylim(0, 1)

# Recall per Fold
plt.subplot(2, 2, 3)
plt.plot(range(1, len(recalls) + 1), recalls, marker='o', linestyle='-',
    color='g')
plt.title('Recall per Fold')
plt.xlabel('Fold')
plt.ylabel('Recall')
plt.ylim(0, 1)

# F1 Score per Fold
plt.subplot(2, 2, 4)
plt.plot(range(1, len(f1_scores) + 1), f1_scores, marker='o', linestyle='-'
    , color='m')
plt.title('F1 Score per Fold')
plt.xlabel('Fold')
```

```
plt.ylabel('F1 Score')
plt.ylim(0, 1)

plt.tight_layout()
plt.show()

# Plot mean ROC curve
plt.figure(figsize=(8, 6))
plt.plot([0, 1], [0, 1], linestyle='--', color='r', label='Random Guessing'
    )
mean_tpr = np.mean(tprs, axis=0)
std_tpr = np.std(tprs, axis=0)  # Standard deviation of TPR
mean_auc = auc(mean_fpr, mean_tpr)
std_auc = np.std(aucs)
plt.plot(mean_fpr, mean_tpr, color='b', label=f'Mean ROC (AUC = {mean_auc
    :.2f}  {std_auc:.2f})')
plt.fill_between(mean_fpr, mean_tpr - std_tpr, mean_tpr + std_tpr, color='
    blue', alpha=0.2)
plt.xlabel('False Positive Rate')
plt.ylabel('True Positive Rate')
plt.title('Mean ROC Curve with AUC and Standard Deviation')
plt.legend(loc='lower right')
plt.show()
```

C.4 Code for Logistic Regression Model Process Pattern included

The import of the necessary libraries are the same as for the pattern excluded code in C.3, hence they won't be shown here. Furthermore, the dataset needs to be changed according to the dataset of the process pattern in the import of the CSV file. In addition, the dataset contains an extra column namely, the process pattern column. Therefore, an additional column is added to the features of the model. Lastly, the same code is used as in C.3 for training the model and plotting the graphs, hence it won't be shown here.

```
#Import CSV file
df = pd.read_csv('../data/PROCESS_PATTERN_DATASET.csv', sep=';')

#Display first 10 rows
df.head(10)

# Define columns of interest
columns_of_interest = [
    'After_Consult by phone',
    'Before_Ultrasound of the abdominal organs.',
    'Before_First outpatient clinic visit.',
    'After_Repeat-pole clinic visit.',
    'Before_Repeat-pole clinic visit.',
    'Before_Consult by phone',
    'Before_Abnormal blood loss',
```

```
    'After_Teleconsult (excluding screen to screen image contact, see
    190019).',
    'Before_First outpatient visit',
    'Before_Repetition',
    'Pattern_Included'
]

# Create X with the specified columns
X = df[columns_of_interest]
# Display the first few rows of X
print(X.head())
y = df['Intervention']
```

D. Folder Structure ZIP File

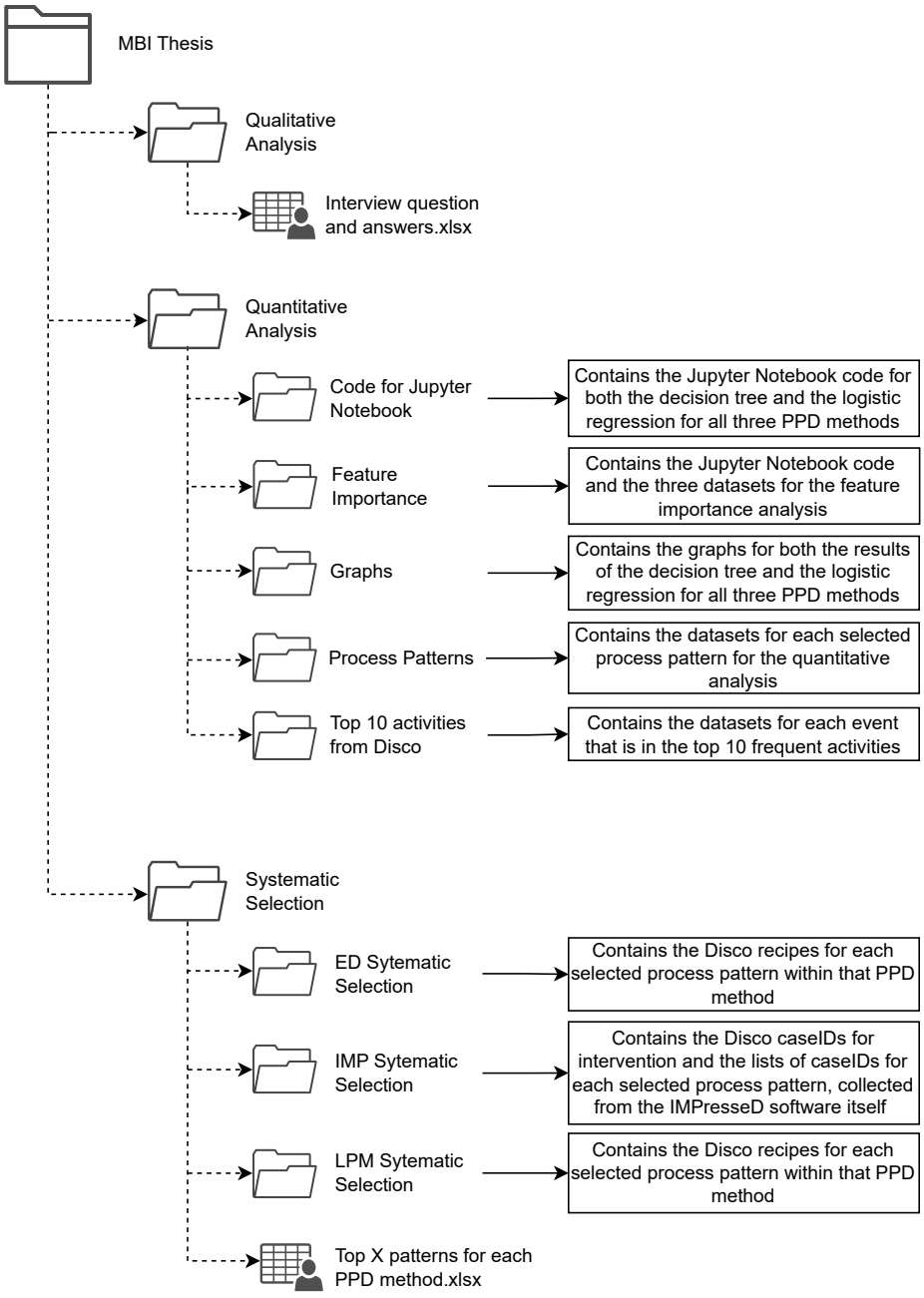

Figure D.1: The folder structure of the ZIP file, with additional content to perform this research