Journal of Biomedical Informatics 61 (2016) 224–236

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Methodological Review

Process mining in healthcare: A literature review

Eric Rojas [a,*], Jorge Munoz-Gama [a], Marcos Sepúlveda [a], Daniel Capurro [b]

[a] Computer Science Department, School of Engineering, Pontificia Universidad Católica de Chile, Chile
[b] Internal Medicine Department, School of Medicine, Pontificia Universidad Católica de Chile, Chile

ARTICLE INFO

Article history:
Received 29 September 2015
Revised 20 April 2016
Accepted 20 April 2016
Available online 22 April 2016

Keywords:
Healthcare
Processes
Process mining
Case studies
Literature review

ABSTRACT

Process Mining focuses on extracting knowledge from data generated and stored in corporate information systems in order to analyze executed processes. In the healthcare domain, process mining has been used in different case studies, with promising results. Accordingly, we have conducted a literature review of the usage of process mining in healthcare. The scope of this review covers 74 papers with associated case studies, all of which were analyzed according to eleven main aspects, including: process and data types; frequently posed questions; process mining techniques, perspectives and tools; methodologies; implementation and analysis strategies; geographical analysis; and medical fields. The most commonly used categories and emerging topics have been identified, as well as future trends, such as enhancing Hospital Information Systems to become process-aware. This review can: (i) provide a useful overview of the current work being undertaken in this field; (ii) help researchers to choose process mining algorithms, techniques, tools, methodologies and approaches for their own applications; and (iii) highlight the use of process mining to improve healthcare processes.

Contents

* Corresponding author.
E-mail addresses: eric.rojas@uc.cl (E. Rojas), jmun@ing.puc.cl (J. Munoz-Gama), marcos@ing.puc.cl (M. Sepúlveda), dcapurro@med.puc.cl (D. Capurro).

http://dx.doi.org/10.1016/j.jbi.2016.04.007
1532-0464/

1. Introduction

The provision of quality hospital services depends on the suitable and efficient execution of processes. Healthcare processes are a series of activities aimed to diagnose, treat and prevent any diseases in order to improve a patient's health. These processes are supported by clinical and non-clinical activities, executed by different types of resources (physicians, nurses, technical specialists, dentists, clerks) and can vary from one organization to another [1]. It is known that healthcare processes are highly dynamic, complex, ad-hoc, and are increasingly multidisciplinary [2], making them interesting to analyze and improve. Healthcare processes improvement might have a high impact on the quality of life of patients. However, improving them is not an easy task and several challenges are always present. There is always the need to reduce the cost of services and improve capabilities to meet the demand, reduce patient's waiting times, improve resources productivity, and increase processes transparency.

The analysis of both clinical and administrative processes can be useful in meeting these objectives, as well as helping to respond any related questions posed by experts. In the past, different strategies have been used to analyze hospital processes, including Business Process Redesign [3], Evidence Based Medicine [4], and Lean [5], among others. This research paper concentrates on the progress made using an emerging discipline, known as Process Mining [6].

Process mining is a relatively young research discipline; [7,8] can be highlighted among the seminal articles. It focuses on extracting knowledge from data generated and stored in the databases of (corporate) information systems in order to build event logs [6]. An event log can be viewed as a set of traces, each containing all the activities executed for a particular process instance. Process-Aware Information Systems (PAIS) [9] are systems that are readily able to produce event logs. Specific examples of such applications include Enterprise Resource Planning systems (e.g., SAP[1]), Customer Relationship Management Systems (e.g., Salesforce[2]), and Hospital Information Systems (e.g., HIS [10]). Event log data are not limited simply to the data from these applications, as many other systems can also provide useful data about process execution. Moreover, data relating to a complex process may come from more than one single source of information.

Fig. 1 shows a general outline of the application of process mining in healthcare. Normally, any activity executed in a hospital by a physician, nurse, technician or any other hospital resource to give care to a patient is stored in a HIS (compound of databases, systems, protocols, events, etc.). Activities are recorded in event logs for support, control and further analysis. Process models are created to specify the order in which different health workers are supposed to perform their activities within a given process, or to analyze critically the process design. Moreover, process models are also used to support the development of HIS, for example, to understand how the information system is expected to support the process execution.

There are three main types of process mining: process discovery, conformance checking, and enhancement. In [6], it is explained how automatic process discovery allows process models to be extracted from an event log; how conformance checking allows monitoring deviations by comparing a given model with the event log; and how enhancement allows extending or improving an existing process model using information about the actual process recorded in the event log.

Having an accurate model of the real behavior of a process improves the capacity of specifying and implementing the process requirements in the HIS that support the process, configuring any additional requirements not included in the system, and supporting the process analysis. In addition, the author in [6] notes the possibility of extending the analysis through other approaches, such as organizational mining, automatic construction of simulation models, model extensions, model repair, predicting process behavior, and recommendations based on history.

Healthcare processes are seen as an area with complex models and which are subject to significant variation over time [2]. Theses variations are caused by multiple factors, including the different conditions of patients and the multiple ways and sequences in which activities can be performed by the resources (physician, nurse, and other healthcare professionals).

The ability to use techniques for discovering process models and analyzing their performance provides valuable opportunities for taking advantage of information stored in HIS event logs. Using process mining techniques in healthcare processes not only ensures such procedures can be firmly understood, but can also generate benefits associated with process efficiency. For example, they can improve the quality of provided services as well as having a positive impact on the management of medical centers.

Besides improving the management of medical centers, additional benefits can be obtained through the application of process mining in healthcare. It can help to identify and understand the real behavior of resources and the patients; to come up with suggestions for redesigning the process; to analyze the performance and to reduce waiting and service times; to obtain insight and improve the collaboration between peers; to predict the behavior of patients according to previous cases; to add additional information to activities such as patient data; to identify which are the activities causing bottlenecks in the process; and can help to identify decision rules applied in different cases.

In order to identify opportunities for applying process mining, it is crucial to be able to understand frequently asked questions posed by healthcare experts regarding such processes. In this study several frequently posed questions were identified in Section 3.3.

The process mining research area has been used in the field of healthcare processes for discovering process models from event logs [11–13], for conformance checking [14,15], and analyzing social networks [12,13,16], among others.

This study aims at identifying and characterizing the case studies where process mining has been applied in the healthcare domain, providing an overview of the state of the art of this field, helping and guiding researchers on what path to follow when applying process mining techniques, methodologies, algorithms and tools; and highlighting some of the advantages of using this discipline.

[1] http://go.sap.com.
[2] http://www.salesforce.com.

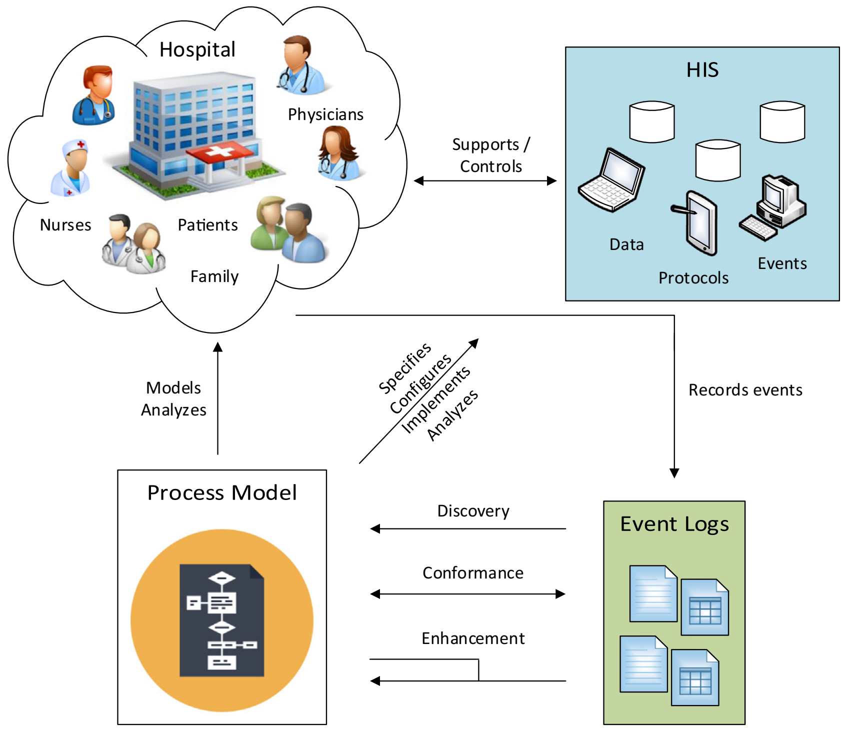

Fig. 1. Process mining in healthcare (based on [6]).

2. Research method

2.1. Background and objectives

Previously, there have been only a few literature reviews about data mining and process mining in the healthcare domain. We have identified some reviews of the application of data mining in several medical fields [17–23]. Regarding process mining in healthcare, there is only a short and specific literature review devoted to case studies related to clinical pathways [24]. However, there is no comprehensive study that collects, characterizes and contextualizes all case studies where process mining has been applied in the healthcare domain.

This review has two main objectives:

Objective 1: Identify and describe existing case studies where process mining has been applied to healthcare processes.

Objective 2: Generate a characterization of existing case studies, including a description of the most important aspects, such as techniques applied, used tools, and data types, among others.

None of the previous objectives has been explored in detail in the past, making this literature review a significant contribution to understanding the overall context of the area and promoting the future application of process mining techniques in the healthcare domain.

2.2. Review questions

A set of questions to guide the review were proposed:

Question 1: Are there any published case studies where process mining techniques and algorithms have been applied in the healthcare domain?

Question 2: What are the main characteristics of the case studies where process mining techniques and algorithms have been applied in the healthcare domain?

Question 3: What are the current results and future trends of the case studies where process mining techniques and algorithms have been applied in the healthcare domain?

2.3. Bibliographic search process

The strategy used for the comprehensive literature search was conducted in three stages. Fig. 2 outlines the process implemented for the search and selection of research material for this review.

The first stage covered the search for papers in three general databases, PubMed, dblp and Google Scholar, using a combination of the keywords "process mining" or "workflow mining", and "healthcare". These keywords were used to try to identify the highest amount of case studies where process mining has been used in healthcare; they were very general so as to include any possible case study. The search was done by the lead author and another of the authors, independently. For the PubMed search the medical expert was the other author involved. Table 1 shows the number of articles found in each database. Notice that in PubMed, as the initial search did not provide any results and given the disciplinary nature of this database, we also carried out a search only using the keywords "process mining" and "workflow mining". In Google Scholar, the search was conducted in the incognito mode, using the option "Search for English results only" and not including patents. Google Scholar only allows access to the first 1000 results of the search, but that seems enough since even in the last 10 pages viewed (901–1000), no relevant items were found.

Afterwards, each reviewer revised the title, abstract, and keywords of each of these papers looking for significant articles about the application of process mining in healthcare. After this initial screening, the articles were reduced to 23 from PubMed, 12 from dblp and 68 from Google Scholar, for a total of 103 papers. After removing the repetitions, 73 were left for the next stage.

The second stage consisted of analyzing articles contained in the case studies repository of the Health Analytics using Process Mining Team from the Department of Mathematics and Computer Science at the Eindhoven University in the Netherlands.[3] The goal of this repository is to provide an overview of all scholarly publications about real-life application of process mining in healthcare, and

[3] http://www.processmining.org/health/papers.

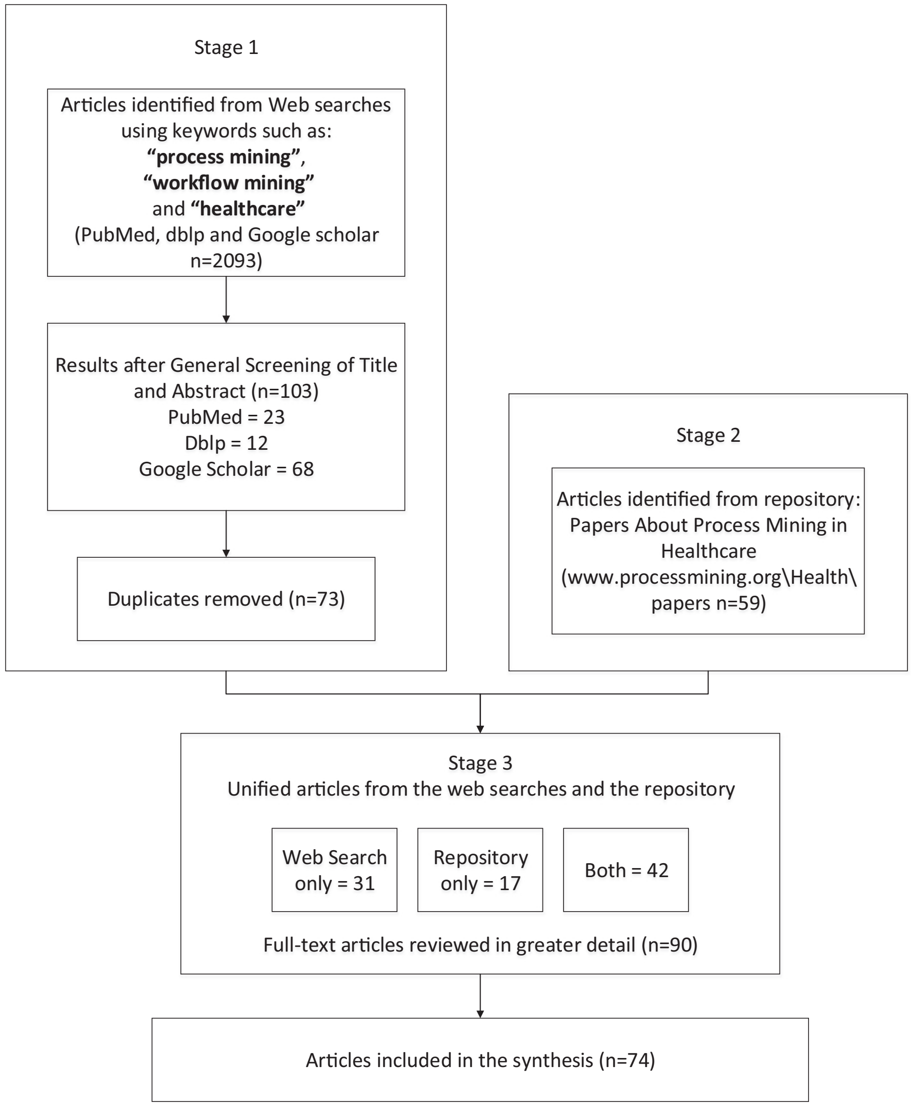

Fig. 2. Search strategy diagram.

Table 1
Initial web search results.

Keywords	PubMed	dblp	Google Scholar
"Process Mining" "Healthcare"	0	13	1670
"Workflow Mining" "Healthcare"	0	0	384
"Process Mining"	26	NA	NA
"Workflow Mining"	0	NA	NA
Totals	26	13	2054
Total = 2093			

contains case studies from different research groups around the world. This stage was executed by the lead author. 59 papers were extracted from the repository.

In the third stage, first the articles from the web searches and the repository were unified. The abstract of each of them was then analyzed by all the authors, applying the inclusion/exclusion criteria and deciding which papers were to be included in the final review. A total of 90 case studies were selected to be analyzed in full detail: 31 from the web searches only, 17 from the repository only, and 42 from both of them. The 90 case studies were reviewed in greater detail by at least two authors, selecting 74 to be included in the synthesis and excluding 16 because they did not include process mining techniques per se.

The source of the selected 74 articles was: 22 from the web searches, 16 from the TU/e repository, and 36 from both of them. They included journal articles, conference presentations, postgraduate and doctoral theses, and a book specific to the area. A complete list of the case studies analyzed in this review can be found in Appendix A. For each one of the reviewed case studies, the most important aspects were identified and analyzed. This included, for example, the methodologies used, the techniques and tools applied, and the medical field from which the data were acquired. Additional aspects could be identified from the case studies, but being this an exploratory analysis, only the main characteristics were included (see Section 6 for Limitations of the study). These main characteristics refer to the common aspects found in the case studies analyzed.

2.4. Inclusion criteria

Following are the inclusion characteristics (IC) used to add a case study to the analysis:

IC 1: All identified articles that include a case study where process mining techniques or algorithms have been applied in the healthcare domain. The degree of analysis is not significant as inclusion criteria.

IC 2: All articles published until February 8, 2016.

IC 3: Only publications in English were included.

2.5. Exclusion criteria

Following are the exclusion characteristics (EC) used to discard a case study from the analysis:

EC 1: Articles that do not include evidence of a case study of process mining in healthcare were excluded.

EC 2: Case studies regarding Clinical Pathways were excluded.

2.6. Quality assessment

To guarantee the quality of the search process a series of activities were undertaken. The initial extraction, analysis and evaluation of the case studies from the web searches in stage one were done manually by the lead author and at least one additional author. The search by keywords was performed in February 2016. All the web searches were done on an incognito mode without cookies or login to guarantee the quality of the results and avoid any impact of previous searches.

In stage three, all authors read the abstracts and verified which article should be included or excluded. Any disagreements with the inclusion/exclusion of a case study were resolved through discussion. In addition to these activities, geographical and medical fields analysis were undertaken to analyze the global representativeness of the case studies analyzed. The geographical analysis of the articles matches with the location of leading research groups in the process mining field around the world; the authors that more often appeared through the different case studies also match with these groups. The medical field could not be used as a quality assessment characteristic because process mining, being an emerging discipline, has not yet been used in all medical fields. However, it is a relevant aspect to analyze. On one hand, the existence of authors who have analyzed various data sets in specific medical fields, allows identifying medical fields where process mining has proved its effectiveness. On the other hand, identifying those medical fields where no case study has yet been reported might encourage other authors to explore the usage of process mining on them in the future.

3. Results

By analyzing the literature addressing the application of process mining in healthcare, a content study was performed. This study identified and classified the relevant common aspects found on those articles. Table 2 outlines a general view and a summary of only the most significant characteristics of these aspects, which are then explained in greater detail below. The complete list of case studies is referenced in each corresponding section.

Sections 3.1 and 3.2 demonstrate the classifications according to process type and data type. Section 3.3 outlines the most frequently questions posed by experts. Section 3.4 identifies the analysis perspectives of process mining. Sections 3.5 and 3.6 outline the tools and techniques used in each case study, while Section 3.7 demonstrates the methodology used for their application. Section 3.8 outlines the implementation strategies of each case study, while Section 3.9 shows the analysis strategies. A geographic classification of the case studies was undertaken in Section 3.10, while an analysis according to the particular medical field in question was conducted in Section 3.11.

3.1. Process types

Process analysts agree about the relevance of being able to identify different types of hospital processes. That way, process mining techniques and algorithms can be applied correctly and appropriately. Four papers of the 74 (5%) specify a classification of processes.

Dumas et al. [9], Refuge et al. [25] and Kaymak et al. [26], specify the existence of two types of hospital processes: medical treatment processes and organizational processes.

The *Medical Treatment Processes* are the clinical processes responsible for managing patients, including tasks ranging from diagnosis to the execution of actions for alleviating each patient.

The *Organizational Processes* focus on the organizational understanding of processes, capturing collaborative information from professionals and their organizational units. For example, this includes assigning tasks by shift and transferring medical information and knowledge between the different types of resources.

Another classification arising from the studies analyzed has been compiled by Mans et al. [1], whereby operational healthcare processes are divided into two types: the first relates to *non-elective care*, including medical emergencies; and the second relates to *elective care*, which includes scheduled standard, routine and non-routine procedures.

3.2. Data types

The types of data determine the processes able to undergo analysis. In addition, the content of these data is used to build the event logs, which allow the process mining techniques to be executed. These event logs relate to the information records containing data from an event associated with an activity, executed through a process, and within a particular case [6]. Only two case studies include classifications of data (3%), which are described as follows:

The first classification is proposed by Kaymak et al. [26], based on data used in their case study. It includes: the *vital signs*, which

Table 2
Main aspects in the reviewed studies.

Aspects	Summary
Process types	*Two classifications:* Medical Treatment Processes and Organizational Processes [9,25,26]; or Non-elective and Elective Care [1]
Data types	*Two classifications:* Vital signs, events, bolus drugs, infusion drugs and inhalation drugs [26]; or Data from administrative systems, clinical support systems, healthcare logistic systems, and medical devices [27]
Frequently posed questions	*Two types:* Specific [27,28] (e.g., do we comply with internal and external guidelines?); or generic [1] (e.g., what happened?)
Process mining perspectives	*Four perspectives:* Control Flow [11–13,16,25], Performance [11–14,29], Conformance [14,15,30,31] and Organizational [11–13,16,32]
Tools	*Three tools:* ProM, Disco and RapidProM
Techniques or algorithms	*Three main techniques or algorithms:* Fuzzy Miner [14,15,33], Heuristics Miner [14,16,34] and Trace Clustering [11,15,35]
Methodologies	*Three main methodologies:* Clustering Methodology [25,35–37]; L* life-cycle Model [38,39] and Ad Hoc Methodologies [40–42]
Implementation strategies	*Three implementation strategies:* Direct [11,13,33,43,44], Semi-Automated [45] and using an Integrated suite [25,46–48]
Analysis strategies	*Three analysis strategies:* Basic, without new implementation [11,16,29,34,49], with new implementation [12,15,34,50] or with analysis from other fields [43,48,50–52]
Geographical analysis	*Two main countries:* The Netherlands and Germany
Medical fields	*Two main medical fields:* Oncology and Surgery

relate to data subject to ongoing monitoring, such as heart rate; the *events*, which are data relating to a series of steps taken in a process, such as the emergency room process; and the *personal data* of patients. This classification also includes data relating to doses of *bolus drugs*, *infusion drugs* and *inhalation drugs*.

The second classification is introduced by Mans et al. [27] and is established in accordance with the data source and its level of abstraction, accuracy, granularity, directness and correctness. The authors, in turn, establish four data sub-types: The first relates to the *data from administrative systems*, which includes data from the administration of accounting services, for example, payment administration software. The second relates to the *data from clinical support systems*, which includes any information system from a particular department in a medical center, based on its specific requirements. For example, this might include software used in pediatrics. The third relates to the *data from healthcare logistic systems*, which includes all data that support the operational process of the medical center, such as the administration of staff shifts. The fourth and final type relates to the *data from medical devices*, such as an X-ray machine.

Ensuring a clear understanding of the available data types helps experts to build event logs and apply process mining in the correct manner.

3.3. Frequently posed questions

Process mining techniques and algorithms provide specialists in medical processes the ability to respond to frequently posed questions about these processes, thereby generating improvement opportunities. According to the frequently asked questions posed by experts, specific data and information should be compiled. Two studies included classifications of frequently asked questions (3%). Mans et al. [27] and Rojas et al. [28] outline five frequently asked questions posed by medical experts: What are the most commonly followed paths and what exceptional paths are followed? Are there any differences between care paths followed by different patient groups? Do we comply with internal and external guidelines? Where are the bottlenecks in the process? And what are the roles and social relationships between medical staff? These questions are updated and established generically in [1] in order to guide the different possible analysis types with process mining:

- What happened?: identifying the need to discover the process executed and its activities (e.g., what is the typical working day of a surgeon?).
- Why did it happen?: understanding the activities and circumstances characterizing the situation/action (e.g., what caused the long waiting list?).
- What will happen?: identifying the circumstances of when or how a specific activity will take place (e.g., is this patient likely to deviate from the normal treatment plan?).
- What is the best that can happen?: identifying possible steps towards specific improvements (e.g., which check should be undertaken first to reduce flow time?).

In summary, these questions represent a guideline in the use and application of process mining techniques, providing correct and precise answers in line with the needs of experts. They can be expanded in greater detail according to the medical specialty and specific needs of a given process.

3.4. Process mining perspectives

According to Van der Aalst [6], there are numerous perspectives in the application of process mining, of which four have been used in healthcare: control flow; performance; conformance; and organizational. All identified case studies have applied at least one of these perspectives.

The control flow perspective, based on discovering the execution order of process activities, has been applied in 60% of the case studies [1,11–14,16,25,26,29,32–35,39–42,45,47,48,50,51,53–75]. The performance perspective, based on analyzing the execution time of activities, identifying bottlenecks, idle time and synchronization time, has been applied in 14% of the case studies [1,11–14,29,39,42,56,59]. Conformance checking, which allows the detection of process deviations in regard to a pre-determined model, has been applied in 21% of the case studies [1,14,15,30,31,38,43,44,46,74–80]. Finally, the organizational perspective, which is based on analyzing the collaboration between resources, has been applied in nine of the 74 case studies, just over 12% [11–13,16,32,44,49,57,81].

3.5. Process mining tools

There are a series of software that enable process mining techniques and algorithms to be applied to an event log in order to generate models, tables and data for analysis.

In healthcare, the most commonly used tool is ProM,[4] which consists of an extensible plug-in open source tool for process mining. ProM has implemented a large number of techniques and algorithms [82] and has been used throughout different case studies (42% or 31 of the 74 case studies) [1,11,12,14,16,29,30,34,37,39,42,47,51,54–57,59,61,62,67,74–76,79–81,83–86].

Additional tools used in case studies include Disco[5] [39,60,63,71,73,80,81,87] used in 8 cases (10%), which consists of a licensed tool with a friendly visual interface for process models and an easy functionality to apply multiple and variable filtering options in event logs, and RapidProM[6] [1], which has a workflow definition functionality based on the plug-ins implemented in ProM and the data analysis solutions implemented in RapidMiner.[7]

3.6. Techniques or algorithms

The main techniques or algorithms used in the case studies are outlined in Table 3. The most commonly used techniques are Heuristics Miner [88], Fuzzy Miner [89] and Trace Clustering [90]. Heuristics Miner is a discovery algorithm that can generate process models and is very robust dealing with noise in event logs [88]. Fuzzy Miner is a configurable discovery algorithm that allows, through its parameters, the generation of multiple models at different levels of detail, helping to deal with unstructured processes [89]. Trace Clustering technique allows the partitioning of the event logs to generate simpler and more structured process models [90].

These are characterized as techniques or algorithms that have already been implemented and which are readily available in open source environments such as ProM. They allow models to be obtained in short execution times and use modeling languages that are easily understood by business analysts (e.g., c-nets [6]). Furthermore, they can help generate models for processes with high variability, such as healthcare processes (e.g., the treatment of different patients with distinct ailments).

Additional techniques used in these case studies include: different algorithms to discover process models (Alpha Miner [6], Genetic Miner [91] and Inductive Miner [92]), verify conformance (Conformance Checker [93]), execute performance analysis (Performance Sequence Analyzer [94]), include expert knowledge

4 http://www.promtools.org.
5 http://fluxicon.com/disco.
6 http://www.rapidprom.org.
7 http://rapidminer.com.

Table 3
Techniques used in the different Case Studies.

Technique	Usage (%)
Heuristics Miner [1,14,16,26,33–35,39,41,42,44,45,49,54,56,59,62,72,86]	26
Fuzzy Miner [1,11,12,14,15,33,37,42,55,56,63,67,73,75,87]	20
Trace Clustering [11,15,35,40,41,44,60,67]	11
Conformance Checker [1,14,40,56,78,80]	8
Performance [11,14,39,42,56,59]	8
Performance Sequence Analyzer [42,44,50,54,84]	7
Ontologies [30,31,73,78]	5
Alpha Miner [16,55,83]	4
LTL Checker [14,44,49]	4
Markov Models [25,53,97]	4
PALIA [46,72,98]	4
Trace Alignment with Tree Guide [12,77,79]	4
Decision Trees [1,80]	3
Dotted Charts [27,84]	3
Genetic Miner [16,54]	3
Parallel Activity-based Log Inference Algorithm Alpha plus Miner [16]	1
Association rule Miner [34]	1
Comp. Miner [42]	1
Cross organizational Comparison [39]	1
Declare Analyzer, Checker, Diagnoser, Replayer and Miner [74]	1
DWS Algorithm [16]	1
Generic Process Model GPM [58]	1
Generic Surgical Process Model (gSPM) [52]	1
Inductive Miner [1]	1
Passage Miner [56]	1
Petri Net Complexity Analysis [75]	1
Prediction Algorithm [99]	1
Replay a log on a Petri Net [75]	1
Rule-Based Property Verificator [40]	1
Semantic Conformance Checker [31]	1
Sequence Clustering [25]	1
Similarity Metric between Two Traces [60]	1
Social Miner [81]	1
Theory of Regions Algorithm [16]	1

(Ontologies [95]) and perform organizational analysis (Social Miner [96]).

3.7. Methodologies

When applying process mining techniques and algorithms, nine papers (12% of the case studies) propose methodologies to be followed. These establish the specific tasks that require completion, for example: obtaining and pre-processing data; the application of techniques and algorithms; and the analysis strategies.

The first methodology uses clustering techniques to conduct analysis. It is based on a fast method for process diagnosis, outlined in [36]. The original methodology executes a diagnosis process through process mining techniques, based on five phases: log preparation; log inspection; control flow analysis; performance analysis; and role analysis. In [25], this methodology was expanded to incorporate a new step after log inspection. This step relates to Sequence Clustering Analysis [25], which incorporates a series of activities for contending with unstructured or spaghetti-type processes. These improvements incorporated activities that helped to discover both typical and infrequent behavior, which were previously undetectable in such complex processes. Subsequently, additional improvements have been made to this particular methodology in [100]. Specifically, these relate to the inclusion of tasks which allow information about the activities executed in a process to be added, as well as event log preprocessing to be undertaken, according to the relevant type of analysis being conducted. The case studies in which this methodology, or similar approaches with clustering, have been applied are [25,35,37].

The second methodology is L* life-cycle Model (Fig. 3). This methodology or model divides a process mining project into five stages: plan and justify; extract; create the control-flow model and connect the event log; create the integrated process model; and provide operational support. In [38], it is used as the suggested methodology for guiding process mining projects, specifically in the Evidence-Based Medical Compliance Cluster (EBMC2). The case studies in which this methodology have been applied are [38,39].

Additional studies which establish or define their own steps or guidelines for applying process mining are outlined in [40–42,71,72,74,81]. None of the main methodologies are domain specific; they can be applied in any other field besides healthcare. The development of methodologies specific for healthcare would be desirable, even for particular medical fields. They could help to better respond to the specific needs of each medical field: making easier to obtain data from information systems, defining specific data reference models for the domain, and proposing how the available tools can be used to answer the frequently posed questions of healthcare experts.

3.8. Implementation strategies

Based on the different case studies uncovered in the healthcare domain, we have compiled a classification of strategies used for implementing process mining.

The first strategy is direct implementation, which consists of applying process mining to a set of data gathered directly from HIS sources for building an event log [1,11,13,33,43,44,49,57,60–62,71,73–75,79,81]. A majority of the case studies use this strategy and consequently obtain models, tables and diagrams for conducting analysis. This strategy poses two main challenges: first, data extraction and building the correct event log; and second, the need to understand the tools, techniques and algorithms available for conducting the analysis.

The second strategy is semi-automated, in which data extraction and the building of the event log are undertaken by custom-made developments [45]. These developments link one or more data sources and extract the data required for building the event log through the use of queries. However, knowledge of process mining tools is still required for applying the available techniques to carry on the correct analysis. This strategy has the disadvantage of being defined in an ad-hoc manner for the extraction of data from specific tools and environments. Only one case study applies this approach.

The third strategy is the implementation of an integrated suite, in which data sources are connected, data are extracted, the event log is built and the implemented process mining techniques are applied. This strategy is applied in 7 case studies (9%) and has the advantage whereby the person using the suite does not require detailed knowledge of how to connect to the data sources or of how the process mining techniques, algorithms and tools work. The great disadvantage of this type of suite is that it has been developed for specific environments and its data sources failing to provide a portable solution. Furthermore, it requires a significant investment in resources in order to conduct the process of implementing the desired algorithms.

Examples of these integrated suites are Medtrix Process Mining Studio [25], Emotiva Tool [46], Careflow Management System Support PM [47], Workflow Management Schemata [48], Business Process Insight Platform [41], Asthma Flow (Prototype Interactive Visual Analytics Tool) [101] and PALIA ILS Suite Web Tool [72].

3.9. Analysis strategies

From the literature review and the case studies identified, the use of three analysis strategies has been established. These strate-

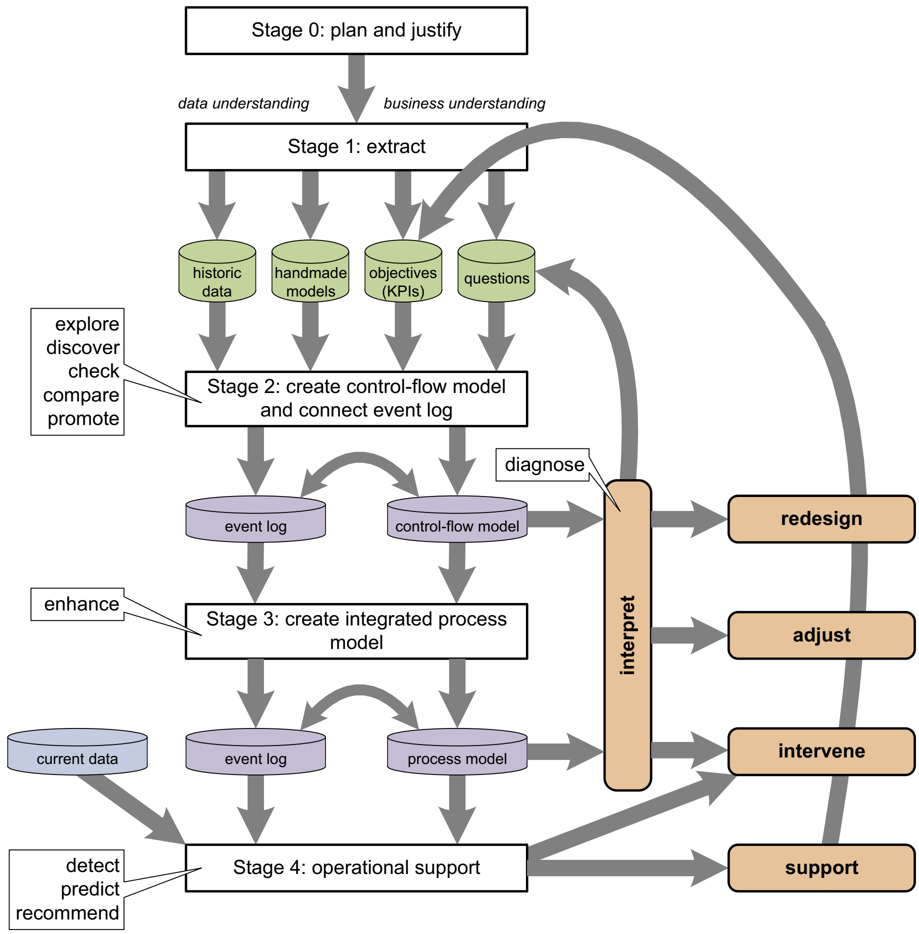

Fig. 3. L* life-cycle model. *Source:* [6].

gies are defined by the way in which they undertake the task of applying process mining techniques and algorithms.

The first is the basic strategy. This consists of taking an event log and applying the process mining techniques and algorithms prevalent in the available tools, without implementing any new technique or algorithm, or applying well-known techniques from other areas, such as statistics. This can be the easiest strategy to perform the analysis and the less resource consuming. It can be performed by medical and process mining experts without requiring the involvement of developers in the project. Examples of case studies in which this strategy has been applied include [1,11,16,27,29,34,49,54,55,71,73,74,81,85] and correspond to 19% of the total cases reviewed.

The second strategy, in addition to applying existing techniques, implements a new process mining technique or algorithm, in certain existing tools, with an objective particular to each specific case study. This requires additional analysis and time to develop the new technique. The main objective with the new implementation is to discover novel ways to deal with processes that are flexible, unstructured and complex, and also to handle large, multidimensional datasets. The general conclusion is that even after the new technique has been applied, the application and analysis through process mining is time-consuming, costly and needs manual effort. There are six case studies (8%), in which this strategy has been applied [12,15,34,50,70,72].

The third and last strategy, in addition to applying techniques and algorithms present in current tools, incorporates analysis of other areas, such as statistical analysis (13%) [12,43,48,50–52,75,79,80,84], analysis that includes OLAP operations [102], such as roll-up or drill-down (1%) [37], data mining (5%) [14,83,84,103], ontologies (3%) [30,31] or simulation models (1%) [55]. This is a noteworthy strategy because it allows new knowledge to be obtained through a combination of existing techniques from different areas. However, it is also more challenging because it requires forming a team of analysts who manage the techniques of each area.

3.10. Geographical analysis

The literature reports on different geographical areas in which case studies of process mining in healthcare have been undertaken. The greatest concentration of case studies is in Europe (around 73%), with a few additional examples in Australia, Asia and North America. There are no records of case studies being undertaken in Africa or Latin America. Specifically, in Europe, The Netherlands is identified as the country with the largest amount of case studies, followed by Germany and Belgium.

3.11. Medical fields

Following the review, it was possible to identify and classify case studies according to the medical field in which data were gathered, from which 22 different areas were recorded. Fig. 4 provides details of the 22 medical fields identified and the number of case studies implemented for each of them. Oncology and Surgery are the medical fields in which the greatest number of case studies were undertaken, with 9 and 8, respectively. The aim of this anal-

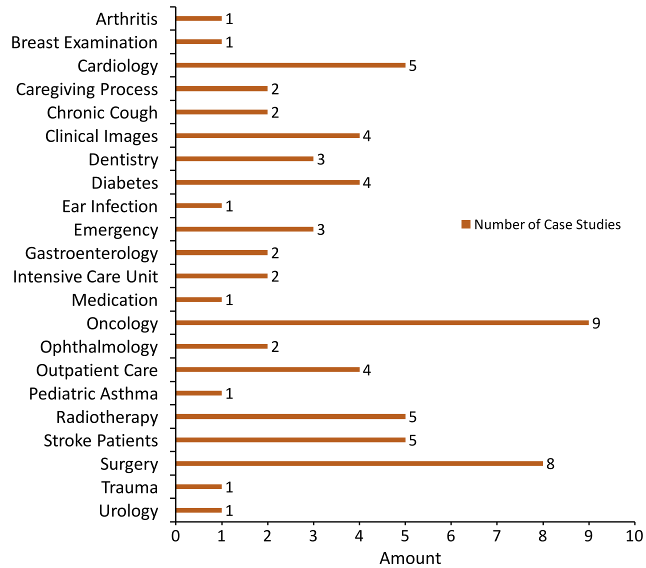

Fig. 4. Case studies by medical fields.

ysis is to show the multidisciplinary character of process mining in healthcare, and its potential application to all medical fields.

4. Discussion

Based on the literature review conducted and the eleven aspects identified, discussion focuses on two approaches: first, identifying the aspect most predominant in the case studies; and second, identifying the main emerging trends.

Regarding the predominant aspects in the application of process mining in healthcare, it is possible to highlight certain key issues. Process mining is primarily applied in the control flow perspective. The most commonly used techniques or algorithms are Trace Clustering, Fuzzy Miner and Heuristics Miner, mainly due to the fact that they adequately manage noise and incompleteness, and allow models to be identified for less-structured processes, which is often the case in healthcare. Furthermore, they allow similar cases to be grouped together, as is the case of trace clustering. ProM is the most commonly used tool, given that it is open source and it is utilized by the leading academic groups in the area. Regarding methodology, there is no predominant approach. It is more common devising an ad hoc methodology for each particular case study. Moreover, the most commonly used implementation strategy is the direct execution of the techniques and algorithms on a manually built event log. It can therefore be identified the need for a standard methodology, which allows medical experts to apply process mining without the direct need of a specialist in the area.

Due to the existence of different HIS architectures (there is almost one for every individual medical center, each with multiple and diverse data sources), the most frequent solution in the use of process mining is the implementation of a specific solution or suite. This does not allow for scalability to other medical centers. Therefore, it is necessary to establish data reference models which help to define standardized data structures in HIS, facilitating the building of event logs.

The countries where the largest amount of case studies have been performed are The Netherlands, Belgium and Germany, because the leading groups of researchers and practitioners in process mining are located in these countries. In turn, the medical field in which process mining has been most widely applied is Oncology.

Seven emerging trends can be identified based on the analysis of the case studies; especially, going over future work sections of the articles reviewed. Below these emerging trends (in no particular order) are described.

First, it should be noted that the most recent articles have tended to implement custom-made solutions for each case study [101]. This has allowed the identification and linking of data sources, the automatic generation of event logs, and the application of process mining techniques and algorithms that have been codified within the respective tool itself. This type of solution allows healthcare experts to apply process mining without the need to understand the technical details, thereby facilitating its adoption.

Second, the establishment of data reference models for HIS, in order they become process-aware information systems. This means new HIS must include information on the activities executed and the order thereof, facilitating the building of event logs and the application of process mining.

Third, the emergence of new tools applied to healthcare, such as RapidProM. These allow the same techniques can be applied to different datasets, facilitating analysis of the same process over different periods of time, or the use of the same strategy for analyzing other processes, which can even belong to other medical centers. This facilitates the increased use of process mining.

Fourth, the increased use of process mining for conformance checking and compliance with specific rules in executed processes. This facilitates the monitoring of processes for verifying compliance with best practices.

Fifth, the interest in benchmarking the performance of one hospital with its pairs, through the comparison of the executed process models. This allows gaps to be detected and best practices and associated standards to be emulated.

Sixth, the inclusion of expertise knowledge in event logs, increasing stored information and facilitating new types of analysis. For example, including medical knowledge through the use of ontologies, for the timely identification of cases requiring special attention. Lastly, an important aspect has been the significant geographical expansion of case studies through time. They now cover more countries, such as the United States.

The importance and growth of the application of process mining in healthcare is causing increasing interest, mainly because it allows making a meaningful usage of stored data. Data can be considered and utilized for analyzing processes, providing real knowledge about their execution and facilitating the identification of improvement opportunities.

5. Future challenges

The future challenges in terms of process mining in healthcare illustrate its growth and importance. The most relevant of these are outlined below.

Despite the significant increase witnessed in the amount of specific developments, including process mining techniques and algorithms for a defined technology environment in a specific medical center, no portable solutions have yet been developed which are capable of being adapted to hospital environments different than those for which they were developed. Future work should concentrate on the implementation of HIS which be process-aware, with the analytical ability to manage data and apply process mining correctly.

One of the negative aspects of current process mining tools is the absence of a good visualization of the process models and the results obtained, especially in complex and less-structured processes, such as those found in the healthcare domain. Improved visualization techniques and visual analytics are required to facilitate the interpretation of the results obtained.

Table 4
Complete list of case studies included in the review.

Author(s)	Title	Ref.
Mans et al.	Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes	[1]
Mans et al.	Application of Process Mining in Healthcare – A Case Study in a Dutch Hospital	[11]
Bose and van der Aalst	Analysis of Patient Treatment Procedures	[12]
Mans et al.	Mining processes in dentistry	[13]
Zhou	Process mining: Acquiring Objective Process Information for Healthcare Process Management with the CRISP-DM Framework	[14]
Kirchner et al.	Embedding Conformance Checking in a Process Intelligence System in Hospital Environments	[15]
Lang et al.	Process Mining for Clinical Workflows: Challenges and Current Limitations	[16]
Rebuge and Ferreira	Business process analysis in healthcare environments: a methodology based on process mining	[25]
Kaymak et al.	On process mining in health care	[26]
Mans et al.	Process Mining in Healthcare: Data Challenges When Answering Frequently Posed Questions	[27]
Mans et al.	Process mining techniques: an application to stroke care	[29]
Grando et al.	Reusing a Declarative Specification to Check the Conformance of Different CIGs	[30]
Grando et al.	Semantic-based conformance checking of computer interpretable medical guidelines	[31]
Mans et al.	A process-oriented methodology for evaluating the impact of IT: a proposal and an application in healthcare	[32]
Kim et al.	Discovery of outpatient care process of a tertiary university hospital using process mining	[33]
Gupta	Workflow and process mining in healthcare	[34]
Caron et al.	Healthcare Analytics: Examining the Diagnosis–treatment Cycle	[35]
Bozkaya et al.	Process diagnostics: a method based on process mining	[36]
Weerdt et al.	Getting a Grasp on Clinical Pathway Data: An Approach Based on Process Mining	[37]
Binder et al.	On Analyzing Process Compliance in Skin Cancer Treatment: An Experience Report from the Evidence-Based Medical Compliance Cluster (EBMC2)	[38]
Partington et al.	Process mining for clinical processes: a comparative analysis of four Australian hospitals	[39]
Caron et al.	A process mining-based investigation of adverse events in care processes	[40]
Lakshmanan et al.	Investigating clinical care pathways correlated with outcomes	[41]
Cho et al.	A Systematic Methodology for Outpatient Process Analysis Based on Process Mining	[42]
Rebuge et al.	A process mining analysis on a virtual electronic patient record system	[43]
Caron et al.	Monitoring care processes in the gynecologic oncology department	[44]
Helmering	Process Mining of Clinical Workflows for Quality and Process Improvement	[45]
Fernández-Llatas et al.	Process mining for individualized behavior modeling using wireless tracking in nursing homes	[46]
Quaglini	Process Mining in Healthcare: A Contribution to Change the Culture of Blame	[47]
Neumuth et al.	Surgical workflow management schemata for cataract procedures	[48]
Caron et al.	Beyond X-Raying a Care-Flow: Adopting Different Focuses on Care-Flow Mining	[49]
Perimal et al.	Gaining insight from patient journey data using a process-oriented analysis approach	[50]
Perimal et al.	Health intelligence: Discovering the process model using process mining by constructing Start-to-End patient journeys	[51]
Neumuth et al.	Analysis of surgical intervention populations using generic surgical process models	[52]
Poelmans et al.	Combining Business Process and Data Discovery Techniques for Analyzing and Improving Integrated Care Pathways	[53]
Fei et al.	Discovering patient care process models from event logs	[54]
Zhou et al.	Process mining based modeling and analysis of workflows in clinical care—a case study in a chicago outpatient clinic	[55]
Suriadi et al.	Measuring patient flow variations: a cross-organisational process mining approach	[56]
Caron et al.	Advanced care-flow mining and analysis	[57]
Peleg et al.	Mining Process Execution and Outcomes – Position Paper	[58]
Montani et al.	Mining and retrieving medical processes to assess the quality of care	[59]
Delias et al.	Supporting Healthcare Management Decisions via Robust Clustering of Event Logs	[60]
Rinner et al.	Cutaneous Melanoma Surveillance by means of Process Mining	[61]
Dagliati et al.	Temporal data mining and process mining techniques to identify cardiovascular risk-associated clinical pathways in Type 2 diabetes patients	[62]
Boere	An analysis and redesign of the ICU weaning process using data analysis and process mining	[63]
Ramos	Healthcare Process Analysis: Validation and Improvements of a Data-based Method using Process Mining and Visual Analytics	[64]
Mans	Workflow support for the healthcare domain	[65]
Maruster and Jorna	From data to knowledge: a method for modeling hospital logistic processes	[66]
Günther et al.	Monitoring deployed application usage with process mining	[67]
Van Genuchten et al.	Is your upgrade worth it? Process mining can tell	[68]
Basole et al.	Understanding variations in pediatric asthma care processes in the emergency department using visual analytics	[69]
Fernandez-Llatas et al.	Diabetes care related process modeling using Process Mining techniques. Lessons learned in the application of Interactive Pattern Recognition: Coping with the Spaghetti Effect	[70]
Micio et al.	RTLS-based process mining: towards an automatic process diagnosis in healthcare	[71]
Fernandez-Llatas et al.	Process Mining Methodology for Health Process Tracking Using Real-Time Indoor Location Systems	[72]
Antonelli and Bruno	Application of Process Mining and Semantic Structuring Towards a Lean Healthcare Network	[73]
Rovani et al.	Declarative process mining in healthcare	[74]
Forsberg et al.	Analyzing PACS Usage Patterns by Means of Process Mining: Steps Toward a More Detailed Workflow Analysis in Radiology	[75]
Dunkl et al.	Assessing Medical Treatment Compliance Based on Formal Process Modeling	[76]
Bouarfa and Dankelman	Workflow mining and outlier detection from clinical activity logs	[77]
Dewandono et al.	Ontology and Process Mining for Diabetic Medical Treatment Sequencing	[78]
Kelleher et al.	Effect of a checklist on advanced trauma life support workflow deviations during trauma resuscitations without pre-arrival notification	[79]
Steeg	Process Mining in Healthcare Mining for Cost and (Near) Incidents	[80]
Rattanavayakorn and Premchaiswadi	Analysis of the social network miner (working together) of physicians	[81]
Riemers	Process Improvement in Healthcare: A Data-Based Method Using a Combination of Process Mining and Visual Analytics	[83]
Staal	Using Process and Data Improving Techniques to Define and Improve Standardization in a Healthcare Workflow Environment	[84]
Paster and Helm	First Steps Towards Process Mining in Distributed Health Information Systems	[85]
Montani et al.	Improving structural medical process comparison by exploiting domain knowledge and mined information	[86]
Overduin	Exploration of the link between the execution of a clinical process and its effectiveness using process mining techniques	[87]
Blum et al.	Workflow mining for visualization and analysis of surgeries	[97]

(continued on next page)

Table 4 (*continued*)

Author(s)	Title	Ref.
Meneu et al.	Heart cycle: facilitating the deployment of advanced care processes	[98]
Van Der Spoel et al.	Process prediction in noisy data sets: a case study in a Dutch hospital	[99]
Doremalen	Process Mining in Healthcare Systems: An Evaluation and Refinement of a Methodology	[100]
Kumar et al.	Exploring Clinical Care Processes Using Visual and Data Analytics: Challenges and Opportunities	[101]
Mcgregor et al.	A process mining driven framework for clinical guideline improvement in critical care	[103]

There is still a great amount of reliance on experts at the moment in which process mining is applied. Efforts should be made to ensure that there are tools or solutions in place which are straightforward to apply, without the need of detailed knowledge of the tools, algorithms or techniques relating to the process mining field. In addition, new methodologies should emerge, which use reference models and be able to consider the most frequently posed questions by healthcare experts.

It is also quite important to incorporate conformance checking in future case studies, in accordance with internal standards or external regulations established by medical institutions. The challenge is to use the latest techniques developed for conformance checking in healthcare. These techniques need first to be improved so as they are capable of working with less structured processes and able to cover a greater number of activities.

Lastly, it is important that benchmarking studies are undertaken between different hospitals, for identifying and emulating success stories.

6. Limitations

It is important to note that the content of this article consists of a methodological review of the current status of process mining in healthcare. It does not go into further detail regarding the results obtained in each case study.

This paper includes an exploratory analysis of the main characteristics of the studies, however, additional aspects could be included and will be considered as future work (data quality, data volume, data issues, ethical and legal constraints, collaboration patterns, etc.).

Searches for this study were undertaken in accordance with criteria specified by the research team (keywords), as well as a general review process of the abstract and the content of the selected articles. There is a possibility that some publications might have been omitted. Results may be subject to limitations of the automated sources used in this work.

7. Conclusions

The application of process mining in healthcare allows health experts to understand the actual execution of processes: discovering process models, checking conformance with medical guidelines, and finding improvement opportunities. This article provides a literature review about the main approaches used to apply process mining in healthcare. It includes a description of process and data types; frequently posed questions; perspectives, tools, techniques and algorithms; methodologies, analysis and implementation strategies; and a breakdown analysis by geographical and medical field. Future challenges and trends have also been identified.

The aim of this review is to serve as a reference guide by including case studies previously conducted in the healthcare domain. It also seeks to highlight the foundations for those aspects that should be taken into account when implementing a process mining project for analyzing healthcare processes.

Author contributions

All authors undertook written and review tasks throughout this study. The process was coordinated by Eric Rojas.

Conflict of interests

No conflicts of interest were reported by the authors regarding this study.

Acknowledgments

The authors would like to thank the Health Analytics Using Process Mining Team from the Department of Mathematics and Computer Science of Eindhoven University in The Netherlands for the repository of case studies compiled, and which has been used in this review paper. This paper was supported by FONDECYT (Chile) grants 1150365, 11130577 and by the Comisión Nacional de Investigación Científica – CONICYT – Ministry of Education, Chile, Ph.D. Student Fellowships.

Appendix A. Details of case studies included in this review

See Table 4.

References

[1] R. Mans, W. van der Aalst, R.J. Vanwersch, Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes, Springer, 2015.
[2] P. Homayounfar, Process mining challenges in hospital information systems, in: Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE, 2012, pp. 1135–1140.
[3] M. Jansen-Vullers, H.A. Reijers, Business process redesign in healthcare: towards a structured approach, Inform. Syst. Oper. Res. 43 (4) (2005) 321–339.
[4] R. Grol, J. Grimshaw, Evidence-based implementation of evidence-based medicine, Joint Commiss. J. Qual. Improve. 25 (10) (1999) 503–513.
[5] Z.J. Radnor, M. Holweg, J. Waring, Lean in healthcare: the unfilled promise?, Soc Sci. Med. 74 (3) (2012) 364–371 (part special issue: Organization studies and the analysis of health systems).
[6] W. Van Der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, Springer Science & Business Media, 2011.
[7] W.M.P. van der Aalst, A.J.M.M. Weijters, Process mining: a research agenda, Comput. Ind. 53 (3) (2004) 231–244.
[8] W.M.P. van der Aalst, B.F. van Dongen, J. Herbst, L. Maruster, G. Schimm, A.J. M.M. Weijters, Workflow mining: a survey of issues and approaches, Data Knowl. Eng. 47 (2) (2003) 237–267.
[9] M. Dumas, W.M.P. van der Aalst, A.H.M. ter Hofstede, Process-Aware Information Systems: Bridging People and Software Through Process Technology, Wiley, 2005.
[10] R. Mans, W.M. van der Aalst, N.C. Russell, P.J. Bakker, A.J. Moleman, Process-aware information system development for the healthcare domain-consistency, reliability, and effectiveness, in: Business Process Management Workshops, Springer, 2009, pp. 635–646.
[11] R.S. Mans, H. Schonenberg, M. Song, W.M.P. van der Aalst, P.J.M. Bakker, Application of process mining in healthcare – a case study in a dutch hospital, in: A.L.N. Fred, J. Filipe, H. Gamboa (Eds.), Biomedical Engineering Systems and Technologies, International Joint Conference, BIOSTEC 2008, Funchal, Madeira, Portugal, January 28–31, 2008, Revised Selected Papers, Communications in Computer and Information Science, vol. 25, Springer, 2008, pp. 425–438.
[12] R.P.J.C. Bose, W.M.P. van der Aalst, Analysis of patient treatment procedures, in: F. Daniel, K. Barkaoui, S. Dustdar (Eds.), Business Process Management Workshops – BPM 2011 International Workshops, Clermont-Ferrand, France,

August 29, 2011, Revised Selected Papers, Part I, Lecture Notes in Business Information Processing, vol. 99, Springer, 2011, pp. 165–166.

[13] R. Mans, H.A. Reijers, M. van Genuchten, D. Wismeijer, Mining processes in dentistry, in: G. Luo, J. Liu, C.C. Yang (Eds.), ACM International Health Informatics Symposium, IHI '12, Miami, FL, USA, January 28–30, 2012, ACM, 2012, pp. 379–388.

[14] J. Zhou, Master: Process Mining: Acquiring Objective Process Information for Healthcare Process Management with the Crisp-dm Framework (Ph.D. thesis, Master's thesis), Eindhoven University of Technology, Eindhoven, 2009.

[15] K. Kirchner, N. Herzberg, A. Rogge-Solti, M. Weske, Embedding conformance checking in a process intelligence system in hospital environments, in: R. Lenz, S. Miksch, M. Peleg, M. Reichert, D. Riaño, A. ten Teije (Eds.), Process Support and Knowledge Representation in Health Care -+ BPM 2012 Joint Workshop, ProHealth 2012/KR4HC 2012, Tallinn, Estonia, September 3, 2012, Revised Selected Papers, Lecture Notes in Computer Science, vol. 7738, Springer, 2012, pp. 126–139.

[16] M. Lang, T. Bürkle, S. Laumann, H. Prokosch, Process mining for clinical workflows: challenges and current limitations, in: S.K. Andersen, G.O. Klein, S. Schulz, J. Aarts (Eds.), eHealth Beyond the Horizon – Get IT There, Proceedings of MIE2008, The XXIst International Congress of the European Federation for Medical Informatics, Göteborg, Sweden, May 25–28, 2008, Studies in Health Technology and Informatics, vol. 136, IOS Press, 2008, pp. 229–234.

[17] M. Marinov, A.S.M. Mosa, I. Yoo, S.A. Boren, Data-mining technologies for diabetes: a systematic review, J. Diabetes Sci. Technol. 5 (6) (2011) 1549–1556.

[18] A. Sharma, V. Mansotra, Emerging applications of data mining for healthcare management-a critical review, in: 2014 International Conference on Computing for Sustainable Global Development (INDIACom), IEEE, 2014, pp. 377–382.

[19] O. Niaksu, J. Skinulyte, H.G. Duhaze, A systematic literature review of data mining applications in healthcare, in: Web Information Systems Engineering—WISE 2013 Workshops, Springer, 2014, pp. 313–324.

[20] I. Yoo, P. Alafaireet, M. Marinov, K. Pena-Hernandez, R. Gopidi, J.-F. Chang, L. Hua, Data mining in healthcare and biomedicine: a survey of the literature, J. Med. Syst. 36 (4) (2012) 2431–2448.

[21] J. Iavindrasana, G. Cohen, A. Depeursinge, H. Müller, R. Meyer, A. Geissbuhler, et al., Clinical data mining: a review, Yearbook Med. Inform. 2009 (2009) 121–133.

[22] P. Ahmad, S. Qamar, S.Q.A. Rizvi, Techniques of data mining in healthcare: a review, Int. J. Comput. Appl. 120 (15) (2015) 38–50.

[23] H.C. Koh, G. Tan, et al., Data mining applications in healthcare, J. Healthcare Inform. Manage. 19 (2) (2011) 64–72.

[24] W. Yang, Q. Su, Process mining for clinical pathway: literature review and future directions, in: 11th International Conference on Service Systems and Service Management (ICSSSM), IEEE, 2014, pp. 1–5.

[25] Á. Rebuge, D.R. Ferreira, Business process analysis in healthcare environments: a methodology based on process mining, Inform. Syst. 37 (2) (2012) 99–116.

[26] U. Kaymak, R. Mans, T. van de Steeg, M. Dierks, On process mining in health care, in: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, SMC 2012, Seoul, Korea (South), October 14–17, 2012, IEEE, 2012, pp. 1859–1864.

[27] R. Mans, W.M.P. van der Aalst, R.J.B. Vanwersch, A.J. Moleman, Process mining in healthcare: data challenges when answering frequently posed questions, in: R. Lenz, S. Miksch, M. Peleg, M. Reichert, D. Riaño, A. ten Teije (Eds.), Process Support and Knowledge Representation in Health Care – BPM 2012 Joint Workshop, ProHealth 2012/KR4HC 2012, Tallinn, Estonia, September 3, 2012, Revised Selected Papers, Lecture Notes in Computer Science, vol. 7738, Springer, 2012, pp. 140–153.

[28] E. Rojas, M. Arias, M. Sepulveda, Clinical processes and its data, what can we do with them?, in: International Conference on Health Informatics (HEALTHINF) 2015, Science and Technology Press (SCITEPRESS), 2015, pp 642–647.

[29] R. Mans, H. Schonenberg, G. Leonardi, S. Panzarasa, A. Cavallini, S. Quaglini, W.M.P. van der Aalst, Process mining techniques: an application to stroke care, in: S.K. Andersen, G.O. Klein, S. Schulz, J. Aarts (Eds.), eHealth Beyond the Horizon – Get IT There, Proceedings of MIE2008, The XXIst International Congress of the European Federation for Medical Informatics, Göteborg, Sweden, May 25–28, 2008, Studies in Health Technology and Informatics, vol. 136, IOS Press, 2008, pp. 573–578.

[30] M.A. Grando, W.M.P. van der Aalst, R. Mans, Reusing a declarative specification to check the conformance of different cigs, in: F. Daniel, K. Barkaoui, S. Dustdar (Eds.), Business Process Management Workshops – BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part II, Lecture Notes in Business Information Processing, vol. 100, Springer, 2011, pp. 188–199.

[31] M. Grando, M. Schonenberg, W.M.P. van der Aalst, Semantic-based conformance checking of computer interpretable medical guidelines, in: Biomedical Engineering Systems and Technologies, Springer, 2013, pp. 285–300.

[32] R. Mans, H.A. Reijers, D. Wismeijer, M. van Genuchten, A process-oriented methodology for evaluating the impact of IT: a proposal and an application in healthcare, Inform. Syst. 38 (8) (2013) 1097–1115.

[33] E. Kim, S. Kim, M. Song, S. Kim, D. Yoo, H. Hwang, S. Yoo, Discovery of outpatient care process of a tertiary university hospital using process mining, Healthcare Inform. Res. 19 (1) (2013) 42–49.

[34] S. Gupta, Workflow and Process Mining in Healthcare (Ph.D. thesis, Master's thesis), Technische Universiteit Eindhoven, 2007.

[35] F. Caron, J. Vanthienen, B. Baesens, Healthcare analytics: examining the diagnosis–treatment cycle, Procedia Technol. 9 (2013) 996–1004.

[36] M. Bozkaya, J. Gabriels, J. Werf, Process diagnostics: a method based on process mining, in: International Conference on Information, Process, and Knowledge Management, 2009. eKNOW'09, IEEE, 2009, pp. 22–27.

[37] J.D. Weerdt, F. Caron, J. Vanthienen, B. Baesens, Getting a grasp on clinical pathway data: an approach based on process mining, in: T. Washio, J. Luo (Eds.), Emerging Trends in Knowledge Discovery and Data Mining – PAKDD 2012. International Workshops: DMHM, GeoDoc, 3Clust, and DSDM, Kuala Lumpur, Malaysia, Revised Selected Papers, Lecture Notes in Computer Science, vol. 7769, Springer, 2012, pp. 22–35.

[38] M. Binder, W. Dorda, G. Duftschmid, R. Dunkl, K.A. Fröschl, W. Gall, W. Grossmann, K. Harmankaya, M. Hronsky, S. Rinderle-Ma, C. Rinner, S. Weber, On analyzing process compliance in skin cancer treatment: an experience report from the evidence-based medical compliance cluster (EBMC2), in: J. Ralyté, X. Franch, S. Brinkkemper, S. Wrycza (Eds.), Advanced Information Systems Engineering – 24th International Conference, CAiSE 2012, Gdansk, Poland, June 25–29, 2012. Proceedings, Lecture Notes in Computer Science, vol. 7328, Springer, 2012, pp. 398–413.

[39] A. Partington, M. Wynn, S. Suriadi, C. Ouyang, J. Karnon, Process mining for clinical processes: a comparative analysis of four Australian hospitals, ACM Trans. Manage. Inform. Syst. (TMIS) 5 (4) (2015) 1–19.

[40] F. Caron, J. Vanthienen, K. Vanhaecht, E. Van Limbergen, J. Deweerdt, B. Baesens, et al., A process mining-based investigation of adverse events in care processes, Health Inform. Manage. J. 43 (1) (2014) 16–25.

[41] G.T. Lakshmanan, S. Rozsnyai, F. Wang, Investigating clinical care pathways correlated with outcomes, in: Business Process Management, Springer, 2013, pp. 323–338.

[42] M. Cho, M. Song, S. Yoo, A systematic methodology for outpatient process analysis based on process mining, in: Asia Pacific Business Process Management, Springer, 2014, pp. 31–42.

[43] Á. Rebuge, L.V. Lapão, A. Freitas, R. Cruz-Correia, A process mining analysis on a virtual electronic patient record system, in: P.P. Rodrigues, M. Pechenizkiy, J. Gama, R. Cruz-Correia, J. Liu, A.J.M. Traina, P.J.F. Lucas, P. Soda (Eds.), Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems, Porto, Portugal, June 20–22, 2013, IEEE Computer Society, 2013, pp. 554–555.

[44] F. Caron, J. Vanthienen, K. Vanhaecht, E. van Limbergen, J.D. Weerdt, B. Baesens, Monitoring care processes in the gynecologic oncology department, Comput. Biol. Med. 44 (2014) 88–96.

[45] P. Helmering, Process mining of clinical workflows for quality and process improvement, in: Healthcare Information and Management Systems Society 2012 Conference (HIMSS2012), Citeseer, 2012, pp. 1–7.

[46] C. Fernández-Llatas, J.-M. Benedi, J.M. García-Gómez, V. Traver, Process mining for individualized behavior modeling using wireless tracking in nursing homes, Sensors 13 (11) (2013) 15434–15451.

[47] S. Quaglini, Process mining in healthcare: a contribution to change the culture of blame, in: D. Ardagna, M. Mecella, J. Yang (Eds.), Business Process Management Workshops, BPM 2008 International Workshops, Milano, Italy, September 1–4, 2008. Revised Papers, Lecture Notes in Business Information Processing, vol. 17, Springer, 2008, pp. 308–311.

[48] T. Neumuth, P. Liebmann, P. Wiedemann, J. Meixensberger, et al., Surgical workflow management schemata for cataract procedures, Methods Inform. Med. 51 (5) (2012) 371–382.

[49] F. Caron, J. Vanthienen, J. De Weerdt, B. Baesens, Beyond X-raying a care-flow: adopting different focuses on care-flow mining, in: Proceedings of the First International Business Process Intelligence Challenge (BPIC'11), 2011, pp. 1–11.

[50] L. Perimal-Lewis, S. Qin, C. Thompson, P. Hakendorf, Gaining insight from patient journey data using a process-oriented analysis approach, Proceedings of the Fifth Australasian Workshop on Health Informatics and Knowledge Management, vol. 129, Australian Computer Society, Inc., 2012, pp. 59–66.

[51] L. Perimal-Lewis, D. De Vries, C.H. Thompson, Health intelligence: discovering the process model using process mining by constructing start-to-end patient journeys, Proceedings of the 7th Australasian Workshop on Health Informatics and Knowledge Management, vol. 153, Australian Computer Society, Inc., 2014, pp. 59–67.

[52] T. Neumuth, P. Jannin, J. Schlomberg, J. Meixensberger, P. Wiedemann, O. Burgert, Analysis of surgical intervention populations using generic surgical process models, Int. J. Comput. Assisted Radiol. Surg. 6 (1) (2011) 59–71.

[53] J. Poelmans, G. Dedene, G. Verheyden, H.V. der Mussele, S. Viaene, E.M.L. Peters, Combining business process and data discovery techniques for analyzing and improving integrated care pathways, in: P. Perner (Ed.), Advances in Data Mining. Applications and Theoretical Aspects, 10th Industrial Conference, ICDM 2010, Berlin, Germany, July 12–14, 2010. Proceedings, Lecture Notes in Computer Science, vol. 6171, Springer, 2010, pp. 505–517.

[54] H. Fei, N. Meskens, et al., Discovering patient care process models from event logs, in: 8th International Conference of Modeling and Simulation, MOSIM 2008, vol. 10, Citeseer, 2008, pp. 10–12.

[55] Z. Zhou, Y. Wang, L. Li, Process mining based modeling and analysis of workflows in clinical care—a case study in a chicago outpatient clinic, in: IEEE 11th International Conference on Networking, Sensing and Control (ICNSC), 2014, IEEE, 2014, pp. 590–595.
[56] S. Suriadi, R.S. Mans, M.T. Wynn, A. Partington, J. Karnon, Measuring patient flow variations: a cross-organisational process mining approach, in: Asia Pacific Business Process Management, Springer, 2014, pp. 43–58.
[57] F. Caron, J. Vanthienen, J. De Weerdt, B. Baesens, Advanced care-flow mining and analysis, in: Business Process Management Workshops 2012, Springer, 2012, pp. 167–168.
[58] M. Peleg, P. Soffer, J. Ghattas, Mining process execution and outcomes – position paper, in: A.H.M. ter Hofstede, B. Benatallah, H. Paik (Eds.), Business Process Management Workshops, BPM 2007 International Workshops, BPI, BPD, CBP, ProHealth, RefMod, semantics4ws, Brisbane, Australia, September 24, 2007, Revised Selected Papers, Lecture Notes in Computer Science, vol. 4928, Springer, 2007, pp. 395–400.
[59] S. Montani, G. Leonardi, S. Quaglini, A. Cavallini, G. Micieli, Mining and retrieving medical processes to assess the quality of care, in: Case-Based Reasoning Research and Development, Springer, 2013, pp. 233–240.
[60] P. Delias, M. Doumpos, E. Grigoroudis, P. Manolitzas, N. Matsatsinis, Supporting healthcare management decisions via robust clustering of event logs, Knowl.-Based Syst. 84 (2015) 203–213.
[61] C. Rinner, R. Dunkl, W. Gall, K.A. Fröschl, W. Grossmann, S. Rinderle-Ma, W. Dorda, H. Kittler, G. Duftschmid, Cutaneous Melanoma Surveillance by means of Process Mining, European Federation for Medical Informatics and IOS Press, 2014.
[62] A. Dagliati, L. Sacchi, C. Cerra, P. Leporati, P. De Cata, L. Chiovato, J. Holmes, R. Bellazzi, et al., Temporal data mining and process mining techniques to identify cardiovascular risk-associated clinical pathways in type 2 diabetes patients, in: 2014 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), IEEE, 2014, pp. 240–243.
[63] J.-J. Boere, An Analysis and Redesign of the ICU Weaning Process using Data Analysis and Process Mining (Ph.D. thesis), Maastricht University Medical Centre, 2013.
[64] L.T. Ramos, Healthcare Process Analysis: Validation and Improvements of a Data-based method using Process Mining and Visual Analytics (Ph.D. thesis, Master's thesis), Eindhoven University of Technology, Eindhoven, 2009.
[65] R. Mans, Workflow Support for the Healthcare Domain (Ph.D. thesis), Technische Universiteit Eindhoven, 2011.
[66] L. Maruster, R.J. Jorna, From data to knowledge: a method for modeling hospital logistic processes, IEEE Trans. Inform. Technol. Biomed. 9 (2) (2005) 248–255.
[67] C. Günther, A. Rozinat, W. van der Aalst, K. van Uden, Monitoring Deployed Application Usage with Process Mining, BPM Center Report BPM-08-11, 2008, pp. 1–8.
[68] M. van Genuchten, R. Mans, H. Reijers, D. Wismeijer, Is your upgrade worth it? Process mining can tell, Software, IEEE 31 (5) (2014) 94–100.
[69] R.C. Basole, M.L. Braunstein, V. Kumar, H. Park, M. Kahng, D.H.P. Chau, A. Tamersoy, D.A. Hirsh, N. Serban, J. Bost, et al., Understanding variations in pediatric asthma care processes in the emergency department using visual analytics, J. Am. Med. Inform. Assoc. 22 (2) (2015) 318–323.
[70] C. Fernandez-Llatas, A. Martinez-Millana, A. Martinez-Romero, J.M. Benedi, V. Traver, Diabetes care related process modelling using process mining techniques. Lessons learned in the application of interactive pattern recognition: coping with the spaghetti effect, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 2127–2130.
[71] R. Micio, F. Fontanili, G. Marques, P. Bomert, M. Lauras, Rtls-based process mining: towards an automatic process diagnosis in healthcare, in: 2015 IEEE International Conference on Automation Science and Engineering (CASE), IEEE, 2015, pp. 1397–1402.
[72] C. Fernandez-Llatas, A. Lizondo, E. Monton, J.-M. Benedi, V. Traver, Process mining methodology for health process tracking using real-time indoor location systems, Sensors 15 (12) (2015) 29821–29840.
[73] D. Antonelli, G. Bruno, Application of process mining and semantic structuring towards a lean healthcare network, in: Risks and Resilience of Collaborative Networks, Springer, 2015, pp. 497–508.
[74] M. Rovani, F.M. Maggi, M. de Leoni, W.M. van der Aalst, Declarative process mining in healthcare, Expert Syst. Appl. 42 (23) (2015) 9236–9251.
[75] D. Forsberg, B. Rosipko, J.L. Sunshine, Analyzing pacs usage patterns by means of process mining: steps toward a more detailed workflow analysis in radiology, J. Digital Imaging 29 (1) (2016) 47–58.
[76] R. Dunkl, K.A. Fröschl, W. Grossmann, S. Rinderle-Ma, Assessing medical treatment compliance based on formal process modeling, in: A. Holzinger, K. Simonic (Eds.), Information Quality in e-Health – 7th Conference of the Workgroup Human–Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2011, Graz, Austria, November 25–26, 2011. Proceedings, Lecture Notes in Computer Science, vol. 7058, 2011, pp. 533–546.
[77] L. Bouarfa, J. Dankelman, Workflow mining and outlier detection from clinical activity logs, J. Biomed. Inform. 45 (6) (2012) 1185–1190.
[78] R.D. Dewandono, R. Fauzan, R. Sarno, M. Sidiq, Ontology and process mining for diabetic medical treatment sequencing, in: Proceedings of The 7th International Conference on Information & Communication Technology and Systems (ICTS), 2013, pp. 171–178.
[79] D.C. Kelleher, R.J.C. Bose, L.J. Waterhouse, E.A. Carter, R.S. Burd, Effect of a checklist on advanced trauma life support workflow deviations during trauma resuscitations without pre-arrival notification, J. Am. College Surg. 218 (3) (2014) 459–466.
[80] S. Van de, Process Mining in Healthcare Mining for Cost and (Near) Incidents (Ph.D. thesis), Eindhoven University of Technology, Eindhoven, 2015.
[81] P. Rattanavayakorn, W. Premchaiswadi, Analysis of the social network miner (working together) of physicians, in: 2015 13th International Conference on ICT and Knowledge Engineering (ICT & Knowledge Engineering 2015), IEEE, 2015, pp. 121–124.
[82] B.F. van Dongen, A.K.A. de Medeiros, H.M.W. Verbeek, A.J.M.M. Weijters, W.M. P. van der Aalst, The prom framework: a new era in process mining tool support, in: G. Ciardo, P. Darondeau (Eds.), Applications and Theory of Petri Nets 2005, 26th International Conference, ICATPN 2005, Miami, USA, June 20–25, 2005, Proceedings, Lecture Notes in Computer Science, vol. 3536, Springer, 2005, pp. 444–454.
[83] P. Riemers, Process Improvement in Healthcare: A Data-based Method using a Combination of Process Mining and Visual Analytics (Ph.D. thesis, Master's thesis), Eindhoven University of Technology, Eindhoven, 2009.
[84] J. Staal, Using Process and Data Improving Techniques to Define and Improve Standardization in a Healthcare Workflow Environment (Ph.D. thesis, Master's thesis), Eindhoven University of Technology, Eindhoven, 2010.
[85] F. Paster, E. Helm, First steps towards process mining in distributed health information systems, J. Electron. Telecommun. 61 (2) (2015) 137–142.
[86] S. Montani, G. Leonardi, S. Quaglini, A. Cavallini, G. Micieli, Improving structural medical process comparison by exploiting domain knowledge and mined information, Artif. Intell. Med. 62 (1) (2014) 33–45.
[87] M. Overduin, Exploration of the Link between the Execution of a Clinical Process and its Effectiveness using Process Mining Techniques (Ph.D. thesis), Eindhoven University of Technology, Eindhoven, 2013.
[88] A. Weijters, W.M. van Der Aalst, A.A. De Medeiros, Process Mining with the Heuristics Miner-algorithm, Technische Universiteit Eindhoven, Tech. Rep. WP, vol. 166, 2006, pp. 1–34.
[89] C.W. Günther, W.M. Van Der Aalst, Fuzzy mining–adaptive process simplification based on multi-perspective metrics, in: Business Process Management, Springer, 2007, pp. 328–343.
[90] M. Song, C.W. Günther, W.M. Van der Aalst, Trace clustering in process mining, in: Business Process Management Workshops, Springer, 2009, pp. 109–120.
[91] A.K.A. de Medeiros, A.J. Weijters, W.M. van der Aalst, Genetic process mining: an experimental evaluation, Data Min. Knowl. Discov. 14 (2) (2007) 245–304.
[92] S.J. Leemans, D. Fahland, W.M. van der Aalst, Discovering block-structured process models from event logs-a constructive approach, in: Application and Theory of Petri Nets and Concurrency, Springer, 2013, pp. 311–329.
[93] A. Rozinat, W.M. van der Aalst, Conformance checking of processes based on monitoring real behavior, Inform. Syst. 33 (1) (2008) 64–95.
[94] W.M. Van der Aalst, B.F. van Dongen, C.W. Günther, A. Rozinat, E. Verbeek, T. Weijters, Prom: the process mining toolkit, BPM (Demos) 489 (2009) 31.
[95] D. Fensel, Ontologies, Springer, 2001.
[96] M. Song, W.M. Van der Aalst, Towards comprehensive support for organizational mining, Decis. Supp. Syst. 46 (1) (2008) 300–317.
[97] T. Blum, N. Padoy, H. Feußner, N. Navab, Workflow mining for visualization and analysis of surgeries, Int. J. Comput. Assisted Radiol. Surg. 3 (5) (2008) 379–386.
[98] T. Meneu, V. Traver, S. Guillen, B. Valdivieso, J. Benedi, C. Fernandez-Llatas, Heart cycle: facilitating the deployment of advanced care processes, in: Proceeding of the 35th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC), IEEE, 2013, pp. 6996–6999.
[99] S. Van Der Spoel, M. Van Keulen, C. Amrit, Process prediction in noisy data sets: a case study in a dutch hospital, in: Data-Driven Process Discovery and Analysis, Springer, 2013, pp. 60–83.
[100] B. van Doremalen, Process Mining in Healthcare Systems: An Evaluation and Refinement of a Methodology (Ph.D. thesis, Master's thesis), Eindhoven University of Technology, Eindhoven, 2012.
[101] V. Kumar, H. Park, R.C. Basole, M. Braunstein, M. Kahng, D.H. Chau, A. Tamersoy, D.A. Hirsh, N. Serban, J. Bost, et al., Exploring clinical care processes using visual and data analytics: challenges and opportunities, in: Knowledge Discovery and Data Mining (KDD): Workshop on Data Science for Social Good, 2014, pp. 1–5.
[102] A. Berson, S.J. Smith, Data Warehousing, Data Mining, and Olap, first ed., McGraw-Hill Inc., New York, NY, USA, 1997.
[103] C. McGregor, C. Catley, A. James, A process mining driven framework for clinical guideline improvement in critical care, Learning from Medical Data Streams 13th Conference on Artificial Intelligence in Medicine (LEMEDS), vol. 765, Citeseer, 2012, pp. 35–50.