Recommendations for enhancing the usability and understandability of process mining in healthcare

Peer-reviewed author version

MARTIN, Niels; De Weerdt, Jochen; Fernández-Llatas, Carlos; Gal, Avigdor; Gatta, Roberto; Ibáñez, Gema; Johnson, Owen; Mannhardt, Felix; Marco-Ruiz, Luis; Mertens , Steven; Munoz-Gama, Jorge; Seoane, Fernando; Vanthienen, Jan; Wynn, Moe Thandar; Boilève, David Baltar; BERGS, Jochen; Joosten-Melis, Mieke; Schretlen, Stijn & Van Acker , Bart (2020) Recommendations for enhancing the usability and understandability of process mining in healthcare. In: Artificial intelligence in medicine, 109 (Art N° 101962).

DOI: 10.1016/j.artmed.2020.101962
Handle: http://hdl.handle.net/1942/32666

Recommendations for enhancing the usability and understandability of process mining in healthcare

Niels Martin[a,b,c,*], Jochen De Weerdt[d], Carlos Fernández-Llatas[e], Avigdor Gal[f], Roberto Gatta[g,h], Gema Ibáñez[e], Owen Johnson[i], Felix Mannhardt[j,k], Luis Marco-Ruiz[l], Steven Mertens[m], Jorge Munoz-Gama[n], Fernando Seoane[o], Jan Vanthienen[d], Moe Thandar Wynn[p], David Baltar Boilève[q], Jochen Bergs[b], Mieke Joosten-Melis[r], Stijn Schretlen[s], Bart Van Acker[r]

[a] *Research Foundation Flanders (FWO), Belgium*
[b] *Hasselt University, Belgium*
[c] *Vrije Universiteit Brussel, Belgium*
[d] *KU Leuven, Belgium*
[e] *Universitat Politècnica de Valencia, Spain*
[f] *Technion - Israel Institute of Technology, Israel*
[g] *Centre Hopitalier Universitaire de Vaudois, Switzerland*
[h] *Università degli Studi di Brescia, Italy*
[i] *Leeds University, United Kingdom*
[j] *SINTEF Digital, Norway*
[k] *NTNU, Norway*
[l] *Nasjonalt Senter for e-helseforskning, Norway*
[m] *Ghent University, Belgium*
[n] *Pontificia Universidad Católica de Chile, Chile*
[o] *Karolinska Institutet, Sweden*
[p] *Queensland University of Technology, Australia*

*Corresponding author

Email addresses: niels.martin@uhasselt.be (Niels Martin), jochen.deweerdt@kuleuven.be (Jochen De Weerdt), cfllatas@itaca.upv.es (Carlos Fernández-Llatas), avigal@ie.technion.ac.il (Avigdor Gal), roberto.gatta.bs@gmail.com (Roberto Gatta), geibsan@itaca.upv.es (Gema Ibáñez), o.a.johnson@leeds.ac.uk (Owen Johnson), felix.mannhardt@sintef.no (Felix Mannhardt), luis.marco.ruiz@ehealthresearch.no (Luis Marco-Ruiz), steven.mertens@ugent.be (Steven Mertens), jmun@ing.puc.cl (Jorge Munoz-Gama), fernando.seoane@ki.se (Fernando Seoane), jan.vanthienen@kuleuven.be (Jan Vanthienen), m.wynn@qut.edu.au (Moe Thandar Wynn), david.marcos.baltar.boileve@sergas.es (David Baltar Boilève), jochen.bergs@uhasselt.be (Jochen Bergs), mieke.joosten-melis@radboudumc.nl (Mieke Joosten-Melis), stijn.schretlen@medtronic.com (Stijn Schretlen), bart.vanacker@radboudumc.nl (Bart Van Acker)

[q]Hospital Universitario Lucus Augusti, Spain
[r]Radboud UMC, The Netherlands
[s]Medtronic Integrated Health Solutions, The Netherlands

Abstract

Healthcare organizations are confronted with challenges including the contention between tightening budgets and increased care needs. In the light of these challenges, they are becoming increasingly aware of the need to improve their processes to ensure quality of care for patients. To identify process improvement opportunities, a thorough process analysis is required, which can be based on real-life process execution data captured by health information systems. Process mining is a research field that focuses on the development of techniques to extract process-related insights from process execution data, providing valuable and previously unknown information to instigate evidence-based process improvement in healthcare. However, despite the potential of process mining, its uptake in healthcare organizations outside case studies in a research context is rather limited. This observation was the starting point for an international brainstorm seminar. Based on the seminar's outcomes and with the ambition to stimulate a more widespread use of process mining in healthcare, this paper formulates recommendations to enhance the usability and understandability of process mining in healthcare. These recommendations are mainly targeted towards process mining researchers and the community to consider when developing a new research agenda for process mining in healthcare. Moreover, a limited number of recommendations are directed towards healthcare organizations and health information systems vendors, when shaping an environment to enable the continuous use of process mining.

Keywords: Process Mining, Healthcare processes, Event log, Process Execution Data, Health Information System, Hospital Information System, Process analysis, Process improvement

1. Introduction

The healthcare sector is confronted with severe challenges, most importantly the contention between tightening budgets and increased care needs due to the aging population [1, 2]. To face these challenges while achieving high quality of care standards, healthcare organizations such as hospitals are becoming increasingly aware of the need to improve their processes (i.e. interrelated sets of activities, decisions and events with a particular goal [3]). Processes play a central role in a healthcare organization's daily operations [1, 4]. They can be subdivided in two categories: clinical processes (e.g. the emergency care process or the treatment process of a particular condition), and administrative/organizational processes (e.g. the inventory management process of materials or the billing process) [5, 6].

Healthcare processes in general, and clinical processes in particular, possess some distinct characteristics compared to common business processes such as the order-to-cash process [6, 7]. Healthcare processes can be characterized as loosely framed and knowledge-intensive [8, 9]. A loosely framed process can be performed in a large, but finite and predefined, number of distinct ways [9]. This relates to the observation that healthcare processes typically exhibit high levels of variation [6]. The knowledge-intensive character implies that the execution of healthcare processes heavily depends on knowledge workers, such as physicians, and the knowledge-intensive decisions they make [8]. These complex decisions are made based on a wide range of criteria, including medical knowledge, patient-related characteristics and the experience of healthcare professionals [5, 6, 8, 10, 11]. Healthcare processes are typically also closely intertwined with each other and are multi-disciplinary, requiring cooperation between clinicians with different expertises, which adds to their complexity [5, 6, 8]. Besides their knowledge-intensive, loosely framed and multi-disciplinary character, healthcare processes are also highly dynamic as they typically continuously change over time due to advances in medical knowledge, technology or administrative procedures [5, 6, 10].

To identify opportunities for process improvement in such healthcare processes, a healthcare organization first needs to gain a profound understanding of the process under consideration. To gather insights in how the process is executed, staff members who are familiar with the process can be brought together for a discussion. This discussion can target the development of process models capturing process insights (such as the order of activities in a clinical process), which forms a basis for a process analysis. However, this is very time- and effort-intensive, and the created process model tends to present an idealized view on the process which might have little connection to reality [12].

To uncover the real behavior of an executed process, a solution can originate from data already collected by health information systems such as a hospital information system. Using the data embedded in the databases of these health information systems, an event log can be generated which contains detailed process execution data for a healthcare process of interest. In this way, the event log contains real-life data about which activities were performed, when they were performed, who performed them and for whom (e.g. for which patient) [12, 13]. Process mining is the research field concerned with the development of techniques to retrieve non-trivial information from such an event log [7, 12]. Over the past decade, the process mining community has developed an extensive set of techniques which convey profound process insights based on real-life data [4, 12]. These techniques relate to the discovery of process models from data, the detection of deviations between an existing model and reality, or the enhancement of an existing process model with, e.g., process performance information [12]. Process mining outcomes can be leveraged to instigate evidence-based process improvement initiatives in healthcare.

The systematic use of process mining in healthcare would be consistent with the *Learning Healthcare System* concept [14] introduced by the Institute of Medicine. One of the basic pillars of a Learning Healthcare System is the implementation of data reuse mechanisms, which allow for learning from data generated during the execution of processes [15, 16]. However, despite the potential of process mining, its uptake in healthcare organizations outside case studies in a research context is rather limited. This observation was the starting point for a two-day international brainstorm seminar[1]. The seminar brought together 18 experts from 11 different countries, both researchers and healthcare practitioners, to reflect upon how to enhance the usability and understandability of process mining in healthcare. This position paper synthesizes the conclusions of the brainstorm seminar. It specifies ten key recommendations that process mining researchers and the community are encouraged to carefully consider when developing a new research agenda for process mining in healthcare. While the focus of our work is predominantly on recommendations for researchers and the research community, three ad-

[1]The international brainstorm seminar took place on June 27th and June 28th, 2019 in Hasselt (Belgium). It was an initiative of the Process-Oriented Data Science for Healthcare Alliance (`https://www.pods4h.com`), supported by the Scientific Research Community on Process Mining (`https://www.srcprocessmining.com`).

ditional recommendations are formulated which explicitly target healthcare organizations and health information systems vendors.

The remainder of this paper is structured as follows. Section 2 provides a a primer to process mining in healthcare, encompassing both an introduction to the topic and an overview of some applications in a healthcare context. Section 3 outlines the key recommendations for process mining researchers and the research community. Section 4 presents the key recommendations for healthcare organizations and health information systems vendors. Section 5 presents a brief conclusion.

2. A process mining in healthcare primer

To provide the required background to appreciate the formulated recommendation, this section provides a primer to process mining in healthcare. Section 2.1 introduces the basic concepts of process mining in healthcare. Building on these concepts, Section 2.2 outlines some applications of process mining in healthcare.

2.1. Introduction to process mining in healthcare

Process mining refers to the retrieval of process-related insights from an event log containing process execution information [12]. Figure 1 positions this field in a broader context. As highlighted in the introduction, the clinical and administrative processes of a healthcare organization are increasingly supported/controlled by health information systems [17]. These systems, such as a hospital information system and a radiology information system, record data about the real-life execution of the process(es) they support in databases. Process execution data included in these databases can be used to create an event log.

An event log consists of a set of events associated to a case such as a patient or a patient visit. Each event represents 'something' that happened within the real-life process which triggers a state transition in a health information system and is registered by the system [18]. An event often relates to the execution of a clinical or non-clinical activity (e.g. the start of a clinical examination for a particular patient or the completion of patient registration), but can also convey that something else took place for a particular case such as the arrival of a message (e.g. test results are available for a particular patient) or the occurrence of an alarm (e.g. a drop in blood pressure of a monitored patient) [3]. An event log minimally contains an ordered set

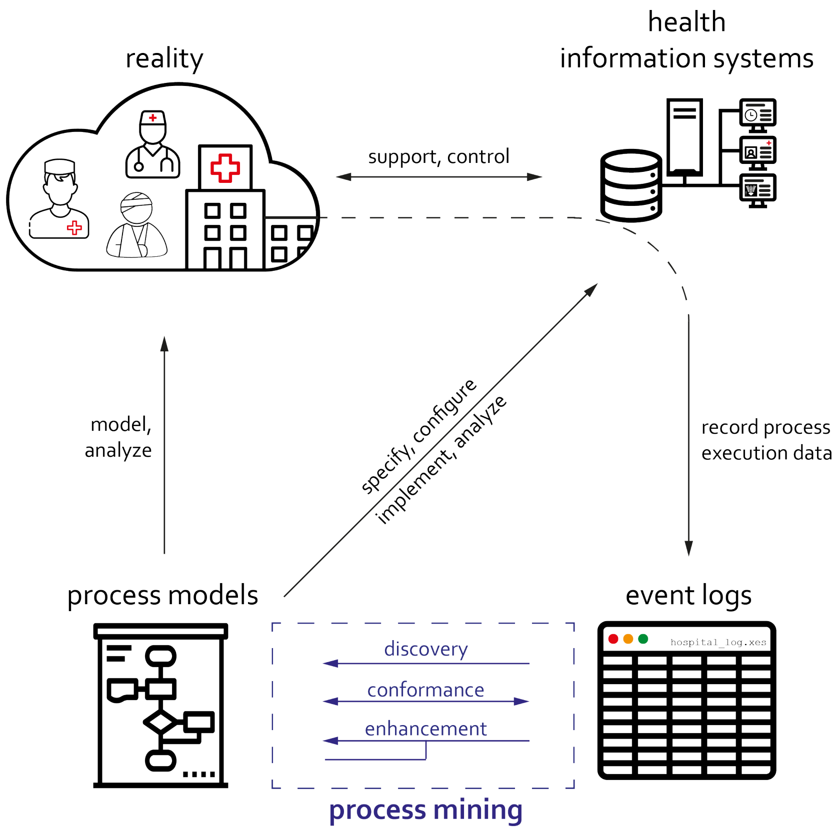

Figure 1: Positioning of process mining (based on [4, 12])

of events for each case, but can also include additional information such as a timestamp expressing when the event took place [12]. Table 1 exemplifies the structure of an event log in a fictitious emergency department setting. For each entry in the event log, i.e. for each event, the following characteristics are recorded [12]:

- **Case identifier (case id):** the (potentially anonymized) identifier of the case (e.g. the patient or the patient visit) to which the event is associated

- **Timestamp:** the time at which the event occurred

- **Activity type:** the label of the activity in the system to which the event is associated

- **Transaction type:** when applicable, the state of the activity type to which the event refers (e.g. its start or its completion)

- **Resource:** the staff member or device associated to the event

For instance: the first row in Table 1 shows that the registration of patient 2478 by administrative clerk Thomas started on June 27th at 10:17:47. He completed the registration at 10:22:02, as shown in the second row of Table 1. While the first two events are clearly related to the execution of 'Register patient', other events reflect, for instance, the availability of blood results for a patient or a blood pressure alarm for another patient. While Table 1, for illustrative purposes, only includes the most common components of an event log, it should be stressed that the event log typically also contains additional attributes about the case or the event. This could include patient attributes (e.g. the patient's age, comorbidities), diagnostic attributes (e.g. blood results, notes from a physician), or care-related attributes (e.g. the medication that was used, whether the admission constitutes a readmission). The availability of such attributes in the event log further enhances the potential usefulness of process mining for evidence-based process improvement.

Taking an event log as an input, process mining techniques generate process-related insights, mainly in the form of process models. In general, three types of process mining are distinguished [12]:

- **Discovery:** Discovery involves the creation of a process model from an event log. An important stream of discovery research relates to

Table 1: Illustration of the event log structure

case id	**timestamp**	**activity type**	**transaction type**	**resource**	**...**
...	...	...	...	...	...
2478	27/06/2019 10:17:47	Register patient	start	Clerk Thomas	...
2478	27/06/2019 10:22:02	Register patient	complete	Clerk Thomas	...
2479	27/06/2019 10:22:18	Register patient	start	Clerk Thomas	...
2472	27/06/2019 10:22:58	Clinical examination	start	Physician Sue	...
2458	27/06/2019 10:23:44	Blood results available	complete	Lab	...
2479	27/06/2019 10:27:58	Register patient	complete	Clerk Thomas	...
2451	27/06/2019 10:29:44	Blood pressure alarm	complete	Device KN0412	...
2451	27/06/2019 10:30:18	Check-up patient	start	Nurse Peter	...
2472	27/06/2019 10:30:27	Clinical examination	complete	Physician Sue	...
2463	27/06/2019 10:31:27	MRI protocol available	complete	Radiologist Sarah	...
2454	27/06/2019 10:32:48	Execute RX-scan	start	Nurse Chris	...
2451	27/06/2019 10:33:04	Check-up patient	complete	Nurse Peter	...
2454	27/06/2019 10:33:45	Execute RX-scan	complete	Nurse Chris	...
...	...	...	...	...	...

control-flow discovery, which retrieves and visualizes the order of activities based on an event log [12]. Within a clinical context, control-flow discovery unearths latent knowledge by making the actual order of clinical activities in the process explicit and visual [2, 19, 20]. To obtain a control-flow model from an event log, a multitude of discovery algorithms have been developed, of which an overview is provided in De Weerdt et al. [21] and Augusto et al. [22]. Specific control-flow discovery algorithms for clinical processes have also been proposed in literature [23, 24]. Despite the predominant research focus on control-flow, some methods also allow for mining, for instance, resource-related models such as social networks [25, 26], resource profiles [27], resource roles [28, 29] and work prioritization patterns [30]. Another example involves the integrated discovery of decision and control-flow models, a technique which has been developed in the context of the clinical process at the emergency department [31].

- **Conformance:** While discovery automatically creates a new model from an event log, conformance checking compares an existing model (e.g. a prescriptive model) with an event log associated to the same process. This allows for both quality assessment of the process model, and the identification of deviations between the model and reality [12]. This latter enables, for instance, judging the compliance of the actual clinical process with clinical guidelines or pathways [32]. Such analyses have been conducted in the context of several pathologies such as unstable angina [7] and colon cancer [33], or in the simulation-based training of ultrasound-guided internal jugular central venous catheter placement [34]. Recent overviews of conformance checking literature are presented in Carmona et al. [35] and Dunzer et al. [36].

- **Enhancement:** Enhancement refers to the extension or improvement of an existing process model based on insights from an event log [12]. This involves, for instance, the extension of a control-flow model with activity durations [37, 38, 39], waiting time information [39], or the decision logic at decision points based on case characteristics [40, 41]. This enables, for instance, the identification of bottlenecks in clinical processes prevailing in reality, which constitute valuable candidates for quality improvement projects [2]. Moreover, the enhanced process model provides a basis to support predictions within a clinical context, e.g. related to the length of stay [42] or the patient recovery time [20].

Within a clinical setting, the medical informatics discipline has traditionally approached process control from a top-down perspective. This implies that experts started from published clinical guidelines and iteratively formalized them to obtain computer-interpretable guidelines [43, 44, 45]. This top-down approach has to cope with limitations such as the complexity to share computer-interpretable guidelines and to adapt them to different local contexts [44]. In contrast, process mining proposes a bottom-up approach which allows the clinician to discover the real process. Moreover, it enables the clinician to discover which interventions would be necessary to reach a better compliance with, for instance, a published clinical guideline. In this way, a process mining approach has important advantages as it does not only inform measures to comply with a specific guideline, but also enables the clinician to investigate which process flows are followed by patients for which no clear guidelines are available. In the latter case, process mining analyses would enhance a healthcare organization's understanding of the patient flow, which can help to improve their treatment processes.

2.2. Applications of process mining in healthcare

For over a decade, process mining has been used to study healthcare processes in a data-driven way. This subsection outlines some applications of process mining in healthcare. Even though it should not be interpreted as an exhaustive list, the applications will illustrate the potential benefits of process mining for healthcare organizations. For a more extensive overview on process mining in healthcare, the reader is referred to literature reviews by Rojas et al. [4], Batista and Solanas [46], and Erdogan and Tarhan [47]. Moreover, dedicated literature reviews have been published on process mining for oncology [48], cardiology [49], primary care [50], and frail elderly care [51].

A key application of process mining involves gaining insights in the order of activities in a healthcare process prevailing in reality (i.e. control-flow discovery). For instance, Kempa-Liehr et al. [20] use process mining to visualize the prevailing pathways for patients suffering from appendicitis to clinicians. Similarly, Kim et al. [52] retrieve the control-flow of the outpatient care process from an event log to gain insights into frequent process paths. These insights can be leveraged to support operational decisions such as the allocation of resources towards activities in such frequently occurring paths. As healthcare processes can exhibit a high degree of complexity and variability, Najjar et al. [11] use clustering techniques to discover treatment processes of elderly patients with heart failure from event logs, which can be used to identify treatment patterns with low mortality rates or patterns that require a closer follow-up. An alternative approach to handle the complexity of healthcare processes is proposed in Chiudinelli et al. [24]. They use topic modeling to synthesize the large number of detailed activities taking place during a particular hospital visit to generate a more high-level overview of the post-surgical treatment processes for breast cancer patients [24]. While the prior illustrations focus on clinical processes in which patients are directly involved, Forsberg et al. [53] study the use of a picture archiving and communication system by radiologists using process mining. This provides opportunities to learn from good practices, as well as becoming aware of potential points of improvement.

Another common use case of process mining in healthcare is understanding how care processes in reality deviate from clinical guidelines or pathways (i.e. conformance checking). Within this context, Huang et al. [7] focus on patient's suffering from unstable angina and propose a method to detect local anomalies in their trajectory compared to the corresponding clinical pathway. Analyzing these patterns can show the need for an update in the clinical pathway or demonstrate that measures to enforce adherence to the pathway are required [7]. Geleijnse et al. [33] study deviations between the clinical pathway and the overall process performance for patients with colon cancer. Vroling [54] relates such deviations from the clinical pathway to patient survival and highlight that deviations in particular treatment phases such as surgery has the largest (negative) impact on survival.

Besides these illustrations, healthcare organizations can use process mining to retrieve valuable insights in their processes in several other ways. They can, for instance, mine social networks to understand interaction patterns between healthcare professionals during the process [55], to identify parts/characteristics of the process contributing to an increased duration of the process [56] and to study the impact on the process of changes such as a move to new facilities [57].

All of the aforementioned applications focused on a single healthcare organization. Process mining can also be used to compare processes between healthcare organizations. In this respect, Mans et al. [58] compare stroke treatment in two different processes. The same context, stroke treatment, is also considered by Montani et al. [59] to demonstrate a general comparison technique for clinical processes. Partington et al. [60] consider patients presenting themselves with acute coronary syndrome symptoms at the emergency department of four Australian hospitals. For this patient population, they compare the order of activities and relevant time intervals between the hospitals [60]. Using process mining for comparison purposes enables cross-hospital benchmarking and learning, creating opportunities to improve the efficiency and quality of care.

While the preceding examples highlight the potential of process mining in healthcare, some applications also cannot reach their full potential. Even though deficient aspects of an application are typically not explicitly articulated in published work, practical experience learns that the availability of reliable data is a key challenge in healthcare [61, 62, 63]. For instance, a process mining analysis at an emergency department focusing on a particular pathology was rescoped because it was impossible to distinguish patients with this condition in the event log. Changes in the project scope due to issues with data availability or quality are common in healthcare [64]. Similarly, the perception among healthcare professionals can also have an impact on a process mining project. For example, a process can typically not be studied from the resource perspective because physicians and nurses fear that the analysis results would be misused for individual performance measurement [63]. These challenges illustrate potential risks for successful process mining projects in healthcare. They will be taken into consideration in the remainder of the paper as the formulated recommendations aim to tackle challenges which are currently present. Taking these recommendations into account in the future will enable healthcare organizations to benefit from the full potential of process mining.

3. Recommendations for process mining researchers and the research community

Despite process mining's great potential to help healthcare organizations understand how their processes are actually executed, its use in healthcare outside a research context is limited. Starting from this observation, a two-day international brainstorm seminar took place to reflect upon how to enhance the usability and understandability of process mining in healthcare. This seminar brought together 18 experts from 11 different countries, both process mining researchers and healthcare practitioners. The aim of the seminar was to formulate recommendations for future efforts and originated from a shared ambition to stimulate a more widespread use of process mining given its potential to support evidence-based process improvement in healthcare. These recommendations are mainly directed to process mining researchers and the research community, but a limited number of recommendations are targeted to healthcare organizations and health information systems vendors.

This section will describe the recommendations for process mining researchers and the research community, which are summarized in Figure 2. While some of these recommendations might also be relevant for other sectors, they also incorporate the specificities of process mining within the healthcare sector. In this way, researchers and the research community are encouraged to take these interconnected recommendations into account when developing a new research agenda for process mining in healthcare.

3.1. Establish a standardized terminology to support process mining in healthcare (RC-1)

Within the healthcare domain, standardized terminologies are common to ensure a shared understanding on the meaning of particular concepts [65]. An example of a widely used terminology is the *International Classification of Diseases* (ICD), developed by the World Health Organization [66]. The most recent revision, ICD-11, defines about 55,000 codes to uniformly refer to injuries, diseases, and death causes worldwide [67]. Another example is the *Logical Observation Identifiers Names and Codes* (LOINC) for laboratory experiments [68]. Also within the context of healthcare processes, terminologies have been developed. Consider, for instance, the emergency department time measures and intervals based on the Performance Measures

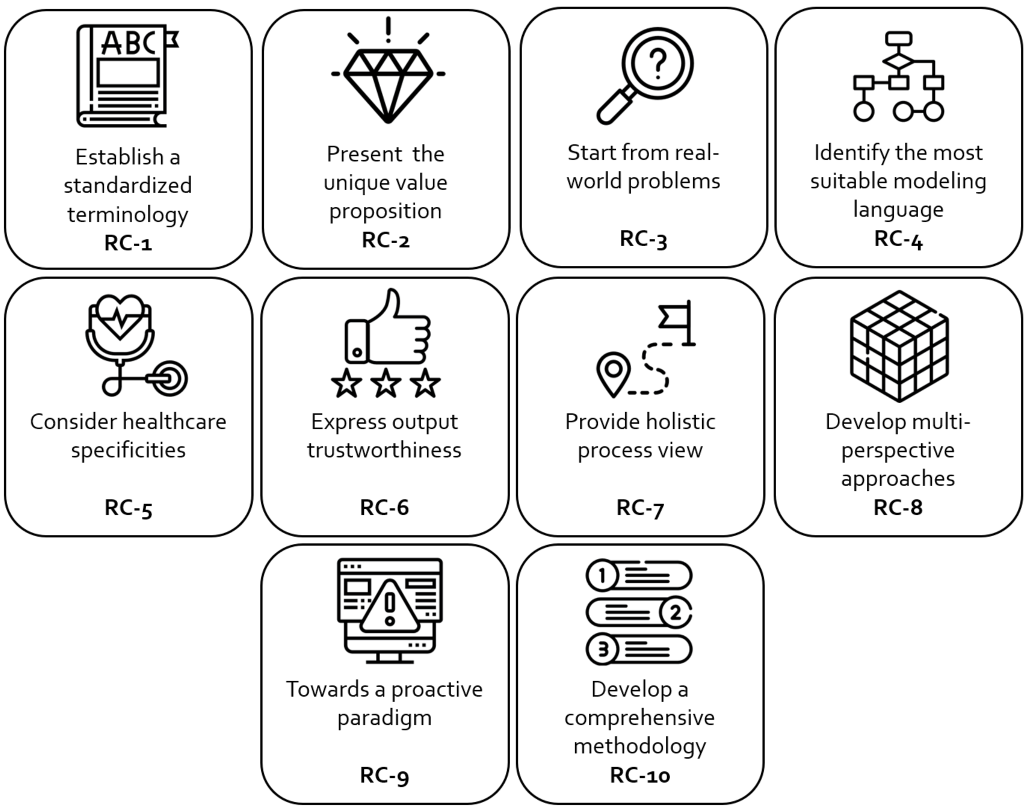

Figure 2: Overview of recommendations for process mining researchers and the research community

and Benchmarking Summit [69, 70]. Some terminologies are designed to act as reference terminologies with a wide coverage of healthcare concepts. That is the case for the *Systematized Nomenclature of Medicine Clinical Terminology* (SNOMED CT), which contains around 350,000 clinical terms [71]. SNOMED CT is based on description logic allowing for the definition of new concepts by combining existing ones [72]. Terminology standardization is also supported by efforts such as OpenEHR [73], which is an architecture aimed to support the interoperability of electronic health record systems. To this end, a wide range of clinical concepts have been standardized in clinical information models called archetypes [74, 75]. OpenEHR archetypes are complementary to other terminology standardization as they can also include references to terminologies such as SNOMED CT [75].

In the process mining field, there is less of a standardization tradition. At the data structure level, a widely recognized process mining standard is the IEEE *Extensible Event Stream* (XES) standard [76]. XES is an XML-based standard providing a format which ensures interoperability in event logs and event streams [77]. This standard also provides a mechanism to add additional extensions [12], such as an extension to express the costs related a particular event [78]. Conversely, at the terminology level, the use of terms used by process miners can be ambiguous and their understanding can depend on the working definitions of individual academics or research groups. This ambiguity occurs in the varying use of terms such as event log, activity log, event, case, trace, activity, resource, classifier, process discovery, and process conformance. For instance: some process mining techniques claim that they use an event log as an input (i.e. each entry represents a single event with a single timestamp), while they in fact require an activity log to start from (i.e. each entry represents an activity instance, which can include multiple timestamps such as the start and completion time). When an event log is available in practice, a conversion to an activity log is required in order to apply the technique. This implies linking corresponding events to each other, an operation which is not always trivial, e.g. when an activity is executed multiple times for a particular patient [79]. The absence of a standardized semantic meaning for process mining terms is especially challenging in a multidisciplinary context, such as healthcare, because process mining terminology will be interspersed with healthcare terminology such as clinical pathway, clinical guideline, length of stay, and evidence-based medicine.

The previous discussion highlights the need to establish a standardized terminology to support process mining in healthcare. This does not imply, by definition, the specification of new terms, but should aim to achieve two key goals. Firstly, the terminology should provide clear descriptions of process mining concepts within the healthcare context. Secondly, it should make, whenever possible, the relationship to existing terminologies in the healthcare domain (such as the procedure branch of SNOMED CT) explicit. The establishment of a standardized terminology requires consensus building within the community to obtain a common understanding. A proposal by a dedicated working group can form a basis for a debate.

Once a standardized terminology is established, it will not only facilitate communication among process mining researchers, but will also enable the use of a uniform terminology towards healthcare professionals. The latter, i.e. having a common language, would greatly enhance the understandability of process mining in the healthcare domain. Moreover, by having an overview of the concepts which are relevant for process mining in healthcare, it can also be determined how the current XES standard (and its extensions) can best be used to support it.

3.2. Present the 'unique value proposition' of process mining in healthcare (RC-2)

The introduction highlighted that, despite the potential of process mining in healthcare, only anecdotal evidence of its systematic use is available until now. To facilitate a more widespread use of process mining in healthcare, process mining needs to be clearly positioned with respect to more established methodologies and systems such as clinical decision support systems [80, 81, 82] and lean management [83, 84]. Consequently, there is a need to present the *'unique value proposition'* of process mining in healthcare, expressing the needs that process mining can fulfill which other methods or systems cannot [85].

A key element in the 'unique value proposition' of process mining in healthcare is its ability to increase transparency by providing actionable and inductive insights into end-to-end processes. Looking at end-to-end processes breaks down departmental silos and potentially even boundaries between healthcare organizations. The actionable insights, derived from real-life process execution data, can be pivotal to improve both clinical and administrative processes. To substantiate this claim, the research community should move beyond individual case studies and develop a taxonomy of problems that process mining can tackle. In this way, a convincing set of key use cases can be composed and communicated. Moreover, for clinical processes, there is a need for research which links interventions based on process mining insights to clinical outcomes. An example of preliminary work in this direction is Vroling [54], where the impact of deviations from the clinical guidelines for colon cancer treatment on patient survival is studied [33, 54]. Similarly, Chiudinelli et al. [24] study the survival probability of breast cancer patients based on their post-surgery treatment process. Relating improvement initiatives inspired by process mining to positive clinical effects will demonstrate that patients can also benefit from process mining. This is an argument that medical doctors are likely to be more susceptible to, compared to a reasoning based on efficiency gains, e.g. due to cost reductions. This, more generally, highlights the need for methods that enable a reliable assessment of the impact of interventions based on process mining outcomes [15].

Developing a substantiated 'unique value proposition' for process mining in healthcare also implies being transparent about the (current) boundaries of process mining. Nevertheless, this is not a plea to position process mining on an island, disconnected from other methods and techniques. In contrast, the close interconnection and complementarity between process mining and other fields such as lean management, predictive analytics, and operations research needs to be stressed. For instance: predictions regarding the expected outcome can be supported by insights in the similarity between the process of a patient who is currently being treated and the trajectory of similar patients in the past [86]. Similarly, simulation is a well-established operations research technique to investigate changes to a process before implementing them using a computer model. It has been extensively applied in healthcare [87]. Process mining can complement simulation by retrieving inductive insights about process behavior, which can be leveraged when building a simulation model [88, 89, 90].

3.3. Start from real-world healthcare problems (RC-3)

To enhance the usability of process mining in healthcare, it is crucial that it tackles real-world problems experienced by healthcare practitioners. This is consistent with one of the key principles of *design science research*, which is a research paradigm that centers around the design, development and scientific study of an artifact which solves a problem [91, 92]. This problem originates from a particular problem context and should be relevant, broadly recognized, and challenging to solve [92, 93]. Within the context of this paper, the problem context refers to (a particular type or set of) healthcare organizations.

In process mining research, healthcare examples are commonly used to show the applicability of a technique that has been developed. However, this does not guarantee that the developed technique solves a real-world problem which is relevant and broadly acknowledged. This is especially attributed to the fact that domain-specific requirements are not explicitly considered during technique development. Researchers active in process mining for healthcare are strongly encouraged to take another perspective. They should stay up to date with the innovation within the healthcare sector. This enables them to identify relevant healthcare problems and treat them as 'first-class citizens'. This requires an important time investment as clarifying the problem typically asks for thorough discussions with healthcare practitioners from different healthcare organizations. Moreover, during the entire research process and when reporting the results, researchers need to reflect upon the implications in healthcare and the actionable insights that can be retrieved from the analyses.

Even when researchers decide to work on a particular clinical process, a researcher needs to know and attribute central importance to the specific problems experienced by clinicians. Starting from these problems, solutions can be developed which leverage process mining to tackle the specific issues that healthcare professionals are confronted with. This also requires awareness of the particularities of the clinical process under consideration as it might influence the design of the techniques. For instance: when the process is composed of a large number of distinct activities, the mining technique might need to incorporate an additional abstraction step. Chiudinelli et al. [24], for instance, propose such an abstraction for the post-surgical breast cancer process as the prime point of interest was the temporal relation between different hospital visits and not the order of activities during one specific hospital visit.

From the previous discussion, it follows that close ties are needed between process mining researchers and healthcare practitioners. This relates to some of the core principles of methodologies such as design science research and action research, which stress the need for a close partnership with the problem context [92]. Moreover, researchers have to be open-minded and should perceive their work as a means to an end, instead of as a goal in itself. Such a mindset is only possible when research starts from real-world healthcare problems.

3.4. Identify the most suitable process modeling language (RC-4)

A key asset of process mining is its ability to retrieve a control-flow model from process execution data. A process model can serve different goals such as creating a shared understanding about the process [94], and supporting the design and configuration of an information system [95]. This recommendation centers around the former, i.e. the use of a process model as a visual instrument to provide healthcare practitioners with insights in their processes. Within the context of a clinical process, the process model can e.g. visualize how the treatment process for a particular pathology takes place in practice, which can be a powerful instrument to evoke process improvement discussions.

To visualize the process flow of a healthcare process, a wide variety of process modeling languages is available. A review by Figl [96] highlights that the used modeling language influences the understandability of the model. Hence, it is important to identify the most suitable process modeling language to represent the output of control-flow discovery algorithms, taking into account the specific healthcare context. In this respect, two key observations that need to be taken into account are (i) the presence of a multitude of process modeling languages available in the business process management domain, and (ii) the presence of modeling language used within the context of clinical guidelines.

The first observation is that, within the business process management field, a large number of different process modeling languages is present. A distinction can be made between procedural, declarative and hybrid languages. In a *procedural control-flow model*, the model shows all possible activity flows, implying that any behavior not shown in the model is not allowed [97]. Hence, a procedural model represents the exact way in which the process can be executed, which typically enhances its understandability for domain experts [98]. Examples of procedural languages are Petri nets/Workflow nets, Yet Another Workflow Language (YAWL), Event-driven Process Chains (EPCs), Business Process Model and Notation (BPMN) [12], and directly-follows graphs [99]. In industry, the ISO-certified standard BPMN (ISO/IEC 19510) has become the *de facto* standard for process modelling [100, 101, 102].

While a procedural approach is appropriate in structured contexts, healthcare processes are typically more variable as, e.g., a clinical process needs to be adjusted to specific patient characteristics [6, 8]. When including all possible execution paths in a procedural process model, this can make the model very complex [98]. In such contexts, *declarative process modelling languages* can be useful. A declarative model consists of a set of constraint that the execution of a process for a patient should satisfy. An example of such a constraint is that 'Execute CT-scan' should be eventually followed by 'CT results available'. Any behavior which satisfies these constraints is allowed [97]. Some declarative process modelling languages, such as Declare or DCR graphs, have a graphical notation, making it possible to use them for communication purposes [103, 104]. However, as the control-flow is represented in an implicit way by means of constraints [105], declarative models are sometimes criticized as being difficult to understand [98]. This especially holds when a declarative approach is used to model a highly structured (part of a) process as it would require a high number of constraints [98].

Recently, researchers recognized that the procedural and declarative process modeling approach should not be treated as mutually exclusive, giving rise to *hybrid process modeling languages* [97, 106]. For instance: BPMN-D is a hybrid language which extends a subset of BPMN (procedural component) with Declare constraints (declarative component). In such a hybrid model, the structured parts of the process are modelled with a procedural language and the variable parts in a declarative way. While hybrid modeling approaches could be valuable within a healthcare context, a recent review by Andaloussi et al. [98] concludes that thorough empirical work on the understandability of hybrid models and their usability for communication purposes is lacking.

The second observation stressing the need to investigate the most suitable way to visualize a healthcare process is that, within the healthcare domain, distinct modeling languages are used to represent clinical guidelines. These modeling languages include Asbru [107], GLIF3 [108], and PROforma [109]. They describe a set of constructs which can be used to model guideline components such as activities that need to be executed and clinical decisions. Some languages also support additional features. For instance, Asbru also enables the incorporation of the guideline's intensions with respect to the process and its outcomes [44]. While these modeling languages are predominantly used to embed clinical guidelines in information systems used by clinicians [44, 110], they should also be considered when looking for a clear communication language as clinicians might be familiar with their notation to a certain extent. For a more extensive overview on modeling languages for clinical guidelines, the reader is referred to the review by Peleg [44].

From these two observations, it follows that a wide range of process modeling languages are available to represent the output of control-flow discovery algorithms to healthcare professionals. Given its impact on model understandability [96], a thorough benchmarking study is required, focusing on differences between languages in terms of the insights that healthcare practitioners retrieve from the resulting models. Besides the main process modeling languages from the business process management domain, modeling languages used for clinical guidelines should also be taken into consideration as healthcare professionals might be more acquainted with them from their clinical practice. Such a comparative study should involve a wide number of healthcare professionals from distinct clinical contexts.

The results of the benchmarking study will provide rich insights in the suitability of existing process modeling languages in particular clinical contexts. This will highlight whether there is a need for a novel process modeling language, potentially leveraging the strengths of several existing languages. At least, the process mining community should carefully assess the similarities and differences between the process model representations used in the process mining field and their counterparts from the clinical domain. This can lead to actionable insights targeting an increase in the understandability of process mining from a representational point of view.

3.5. Take into account healthcare specificities during technique development (RC-5)

Besides tackling real-world healthcare problems (Section 3.3), process mining researchers also need to take healthcare specificities into account during technique development. Consider, for instance, the fact that the flow of a clinical process can be heavily influenced by patient conditions which might not be directly related to the process under analysis. When techniques are tailored to the specific context of healthcare, this will facilitate their use in healthcare practice. For illustrative purposes, the following two healthcare specificities will be discussed: (i) the robustness for temporal changes, and (ii) the ability to focus on infrequent behavior. Many of the existing process discovery and conformance checking techniques are not suitable for analyzing healthcare processes exhibiting such behaviors.

Clinical processes frequently change over time due to a variety of factors such as advances in medicine and innovations of medical equipment. Depending on the type of change, healthcare professionals might be aware that a change happened at a particular point in time. Even when healthcare professionals know that a change took place, they might not have insights in the effects on the process or the (potentially unintended) side-effects. Process mining researchers need to be aware of the existence of temporal changes and techniques should be robust for such changes. The latter especially holds when techniques are intended to be used on event logs with a longer time horizon. Research on the notion of concept drift, which is the term used to express process changes, can be a valuable starting point [111, 112]. However, compared to other contexts, where processes remain stable for extended periods of time, temporal changes should receive more attention in a clinical setting.

Another example of a healthcare specificity is the need to focus on infrequent behavior. In many other application domains, infrequent behavior (such as process execution paths occurring very rarely) is not considered informative as it might distract domain experts from the main process behavior. Hence, some existing process mining algorithms filter out such infrequent behavior to obtain a more understandable process model. For instance: the premise of the heuristics miner (a control-flow discovery algorithm) is that infrequent behavior should not be included in a control-flow model [12, 113]. Depending on the purpose of the analysis, infrequent behavior could be of utmost interest to clinicians as it can demonstrate the need to, e.g., change or sharpen medical protocols. Hence, process mining techniques should have the ability to highlight and analyze infrequent paths to retrieve valuable insights from them. In this respect, efforts such as Hompes et al. [114] and Mannhardt et al. [115], can be a starting point. Hompes et al. [114] use cluster analysis to discover frequent process variants, as well as exceptional paths, in a cancer treatment process. Mannhardt et al. [115] present an approach to distinguish infrequent paths from random noise using data attributes.

The aforementioned examples highlight the importance of tailoring process mining techniques to the specific needs of healthcare. In this respect, interactive process mining methods are promising as they enable healthcare experts to leverage specific domain expertise to obtain useful process mining outcomes within a particular healthcare context. Some interactive approaches have already been presented, for instance to investigate the process of stroke patients at the emergency department [116], or to interactively explore healthcare processes, which is demonstrated for the diabetes treatment process [117]. Moreover, a preliminary method to interactively perform event log cleaning within a healthcare context has recently been proposed in Martin et al. [118]. Additional pointers in the general process mining literature relate, for example, to the inclusion of domain knowledge during control-flow discovery [119, 120].

3.6. *Express the trustworthiness of process mining output (RC-6)*

Since the early 1990s, evidence-based medicine reflects a paradigm in which decision-making in medicine is systematically improved by considering scientific evidence [121]. In evidence-based medicine, different degrees of trustworthiness are attributed to different types of scientific evidence. For instance, the results of a randomized controlled trial are considered as more trustworthy than outcomes of an observational study [122]. In turn, the results of a meta-analysis are considered more trustworthy than the outcomes of a single randomized controlled trial [123].

In an effort to obtain a more widespread use of process mining by healthcare practitioners, it has to meet the key standards which are applicable in the medical domain. Hence, it is important that the output of process mining techniques expresses its trustworthiness. While actions such as showing the frequency of arcs in process models and reporting the conformance metrics are valuable, a more sophisticated way to transparently express the trustworthiness of models is warranted. A starting point is identifying the key elements which influence the trustworthiness of a particular process mining result such as the number of care episodes which are included, and the data sources that have been used.

An aspect that will also greatly influence the trustworthiness of process mining output is the data quality of the event log. Consistent with the *'garbage in - garbage out'* principle, the quality of all process mining analyses ultimately depends on the quality of its input data [2]. This is highly relevant in a healthcare context, where it is not always possible to extract high-quality data from health information systems [124, 125]. A case study at the Maastricht University Medical Centre showed that the three most frequently occurring data quality issues in health information systems data are missing events (i.e. events that took place, but were not logged), imprecise timestamps (i.e. timestamps recorded at the day level), and imprecise resource information (i.e. resource information not referring to a specific staff member) [2]. Another common issue in event logs is that healthcare professionals record their actions in the health information system at a later point in time, and potentially even in batch for several patients [126]. This is highly problematic for process mining as it implies that the timestamps in the event log no longer correspond to the time at which the activity has actually took place. Besides, for instance, biasing insights in the duration of activities, it can also falsify the order of activities when activities are recorded in a different order than they were executed.

In recent years, there has been an uptake in research on data quality within the process mining field. Literature focuses on the identification of data quality issues [2, 126, 127, 61, 128], and the mitigation of one or more of these issues using a particular heuristic [61, 128, 129, 62, 130, 131, 132]. Besides the need to thoroughly assess data quality [126] and keeping a structured data quality register [133], data quality issues should also be reflected in process mining outputs. It could, for instance, be expressed that the trustworthiness of particular aspects of the output (e.g. particular connections between activities in a control-flow model) is lower due to data quality problems.

3.7. Provide a holistic process view (RC-7)

To anchor process mining in a sustainable way in the healthcare sector, it needs to be able to provide healthcare practitioners with a holistic process view. This can be achieved by leveraging all sources of process-related data which are available. Traditionally, process-related data originates from, for instance, a hospital information system. Besides structured fields, many entries in a hospital information system still consist of free text fields, such as the notes of clinicians [134]. Important data about the execution of the process, such as case/event attributes, can be embedded in such free text fields. To unearth insights from unstructured textual data to consider it for process mining purposes, research efforts at the boundary between natural language processing and process mining are required [135, 136]. Within the context of healthcare processes, efforts such as Najafabadipour et al. [134], proposing a method to retrieve the trajectory of lung cancer patients from textual clinical notes, are promising. Epure et al. [137] consider the context of written online conversations about health-related topics to automatically identify speech intentions such as complaining or disagreeing. Gathering insights from unstructured textual data about the process will become increasingly relevant when technologies such as conversational agents will be more intensively used in clinical processes [138, 139].

Besides traditional sources of process-related data, such as a hospital information system, technological advances will make other sources of data available. One notable evolution is the *Internet of Things* (IoT). IoT refers to a network of interconnected devices, equipped with tags or sensors, which enables continuous monitoring of a particular phenomenon [140, 141]. The number of IoT-applications in healthcare is rapidly increasing, as demonstrated by recent reviews such as Nguyen et al. [142] and Ahmadi et al. [143]. Examples of IoT-applications are systems to monitor patients suffering from a chronic disease from home, or an intelligent pill box to assist the elderly with their medication [142]. Another example is FibriCheck, a smartphone application which uses the phone's built-in camera to monitor the heart rate of patients and to detect atrial fibrillation [144]. Due to the increasing presence of tags or sensors in a multitude of connected devices, new sources of process-related data become available. This provides major opportunities to provide a more holistic view on the process as, e.g., treatments which are conducted at home can also be taken into consideration.

Another dimension of a holistic process view involves moving beyond the boundaries of a single healthcare organization (and the data from its information systems). Within the realm of integrated care, patients should experience a continuum of care instead of fragmented care [145, 146]. This implies that a care process is not confined to the hospital's walls, but also encompasses other care entities such as primary care and home health nursing. For instance: a cancer patient might visit a specialized hospital for surgery, have regular check-ups at a local hospital, and take some therapies at home supported by home health nurses. Evidently, this will pose challenges as process execution data will not only be spread over different databases of a single institution, but over a wide number of databases in several institutions. Moreover, healthcare organizations might be reluctant to share data. In this respect, blockchain is a promising technology as it facilitates data exchange among healthcare organizations [147, 148, 149]. Initial approaches on how to use blockchain data for control-flow discovery purposes have recently been proposed [150]. Great care should also be taken to protect the patient's privacy in response to international privacy regulations. Solutions can originate from privacy preservation techniques that allow distributed computations without extracting data from the healthcare organizations where it has been recorded [151]. Despite such challenges, the cross-organizational perspective needs to be taken into consideration when moving towards holistic process mining in healthcare.

3.8. Need for multi-perspective approaches (RC-8)

To enhance the usability of process mining in healthcare, multi-perspective approaches are required. In process mining research, a multi-perspective approach connects several perspectives on the process, such as the control-flow, resource or time perspective [152]. This implies that a process model will not only depict the control-flow, but also, amongst others, includes information about the healthcare professional (or team of professionals) involved, and the waiting times in several stages of the process.

Besides this angle on multi-perspectivism, process mining in healthcare should also have the ability to analyze a process from the perspective of different stakeholders. These stakeholders include physicians, nurses, hospital management, and the patient. Different stakeholders can have different information needs. For instance: while hospital management might be interested in the dominant process flow, physicians might be especially interested in infrequent behavior and its clinical impact.

To date, process mining research attributed limited explicit attention to the patient perspective, i.e. to identify and study the patient's journey [153]. However, this constitutes a valuable direction for research as studying healthcare processes from the patient's perspective will help healthcare organizations to move towards more patient-centered care, which is considered as one of the most important determinants for care quality [154]. Considering the patient perspective can be challenging as it, for instance, involves moving beyond the boundaries of a single healthcare organization (as outlined in the previous recommendation).

3.9. Towards a proactive paradigm (RC-9)

Process mining mainly focuses on providing insights in a process using historical process execution data. Complementary to this retrospective paradigm, healthcare would also greatly benefit from a proactive paradigm. In a proactive setting, process mining techniques generate actionable insights which can be brought into practice while the process is still running. This relates to the notion of process mining providing online operational support [155]. Proactive process mining can enhance the use of process mining in healthcare as it directly supports the daily operations of healthcare practitioners. For instance: process mining can add a context- and process-aware dimension to decision support systems in healthcare [156]. In this way, process mining can inform taking proactive actions when similarities are detected between treated patients or historic process states, and the current patient/state. This requires process similarity measures, where works such as Combi et al. [86] within the context of stroke can be leveraged. Proactive applications can have a trickle-down effect on the use of other techniques once healthcare practitioners are convinced of the added value of process mining.

Research on the proactive use of process mining should be closely linked to existing work on predictive analytics, where the focus is on the prediction of, e.g., the onset or progression of diseases [e.g. 157, 158] or the effectiveness of a particular therapy [e.g. 159].

To further enhance their usability, proactive process mining techniques can be extended with the ability to learn from user interaction. Consider, for instance, a system which automatically raises alerts to a physician when a clinical guideline is not followed. When a particular type of alert is systematically discarded by the physician, the system could autonomously change the alert threshold or recommend the user to alter a particular setting. In this respect, principles from active learning could be leveraged [160].

3.10. Develop a comprehensive methodology for process mining in healthcare (RC-10)

Besides complying with the aforementioned recommendations when, e.g., developing a new process mining technique, the usability of process mining would also be enhanced by guiding healthcare organizations regarding its use. In this respect, there is a need for a comprehensive methodology for process mining in healthcare. Its comprehensive character relates to the fact that it should range from the identification of process-related questions or challenges, over the collection and preparation of process execution data, until the application of suitable process mining techniques to answer the questions or handle the challenges.

Within the process mining field, the L*-methodology [12], the Process Diagnostics Method [161] and the PM^2-methodology [162] provide a high-level overview of the steps that need to be taken when a process mining analysis is performed. However, to provide more targeted support to healthcare organizations, a methodology which takes into account and anticipates upon the specificities of healthcare and clinical processes is warranted. In this respect, a starting point for the development of the methodology can be a synthesis of key use cases of process mining in healthcare. These use cases can be used to group existing techniques and to add pointers to reliable implementations which are available. The latter highlights the need to have well-documented implementations generating clear outputs. Other questions that should receive attention in the methodology include: How to construct an event log which actually captures all relevant dimensions of a healthcare process? How to handle unstructured data such as a clinician's notes in the patient's file? How to ensure that the process mining analysis provides the information that a clinician actually needs to tackle the real problems they face in daily practice? Regarding the latter, existing efforts relating to the use of interactive approaches [116, 118, 120, 163], the use of a question-driven methodology [164], or the (Simplified) Clinical Pathway Analysis Method [165, 166] constitute potential starting points.

When a comprehensive methodology is in place, there will be support for healthcare organizations and clinicians during their process mining endeavors. The presence of such an instrument might also provide the required reassurance for healthcare organizations who are considering the adoption of process mining. In this way, the methodology will both enhance the usability of process mining in healthcare, and will ensure that healthcare organizations get the full potential out of the available techniques. Moreover, by providing an end-to-end view, the methodology will also highlight research challenges for the process mining community in the form of stages which are currently inadequately supported from an algorithmic or methodical point of view.

3.11. Summary

From the previous, it follows that process mining researchers and the research community play an important role in enhancing the usability and understandability of process mining in healthcare. While many recommendations are situated, at least for an important part, at the technical level (RC-4, RC-6, RC-7, RC-8, RC9), efforts at a non-technical level will also prove indispensable to obtain a more widespread use of process mining. Such non-technical efforts relate to clearly communicating the value of process mining for healthcare practitioners using a standardized terminology (RC-1, RC-2), and providing the required support to healthcare organizations to use these techniques (RC-10). Moreover, researchers should have a mindset in which close collaboration with healthcare organizations is a natural reflex. Such a mindset ensures that future process mining techniques explicitly incorporate healthcare specificities and are able to tackle real-world healthcare problems (RC-3, RC-5).

4. Recommendations for healthcare organizations and health information systems vendors

The predominant focus of this paper is on providing recommendations to process mining researchers and the research community to enhance the usability and understandability of process mining in healthcare. However, healthcare organizations and health information systems vendors also play a key role in a more widespread use of process mining as they should provide an environment which enables a continuous use of process mining. To this end, three key recommendations are provided to shape such an environment. The remainder of this section will outline these recommendations, which are summarized in Figure 3.

Figure 3: Overview of recommendations for healthcare organizations and health information systems vendors

4.1. Invest in training opportunities for healthcare professionals (HOV-1)

Currently, healthcare organizations tend to make use of external expertise when conducting a process mining project, e.g. by entering in a partnership with a university or a consulting company. However, to embed process mining in a sustainable way in healthcare organizations, it is important that internal expertise is also built up. Besides some dedicated staff members with a more extensive process mining background, a minimal level of data and process literacy at the level of healthcare professionals or hospital administrators is also required. The latter is, amongst others, important to formulate the right questions that process mining should tackle and to understand the analysis results. To this end, healthcare organizations will need to offer training opportunities to promote healthcare analytics in general and process mining in particular. During such training sessions, clinicians should be motivated to identify process-related challenges they are confronted with in their clinical practice, as these constitute starting points for using process mining.

Enhanced training will ensure that healthcare organizations will treat the available data in their health information systems as a strategic asset. This will also highlight the need to retain full control over their data, even when the maintenance of their systems is outsourced [167]. Moreover, training can help healthcare professionals to see data-driven process analysis as an opportunity to improve patient care, instead of as a threat solely focused on cost reduction or individual performance measurement.

The process mining community can be a facilitator by providing training materials for healthcare professionals in a variety of forms, including dedicated training programs, online courses[2] or hands-on textbooks. Moreover, as was already pointed at in Section 3.10, process mining researchers have to make sure that the techniques they develop are publicly available, deliver clear outputs, and are well-documented.

4.2. Promote efforts to improve the quality of recorded data (HOV-2)

As highlighted earlier in this paper, process execution data from health information systems often suffers from data quality problems [2, 124, 125, 126]. Data quality issues, such as missing events or incorrect timestamps, impede many of the existing process mining techniques to reach their full potential. While research is being performed on the improvement of data quality using heuristics [e.g. 61, 62], this will always remain suboptimal compared to more accurate data recording at the source.

To enhance the potential of healthcare organizations to benefit from process mining, efforts to improve the quality of recorded process execution data are warranted. These efforts can be situated at two levels: (i) at the health information system level, and (ii) at the level of logging attitudes and mechanisms.

At the level of health information systems, vendors could increase the process-aware character of their systems. This would cause the data to be recorded in a more process-oriented way, facilitating its use within a process mining context. Moreover, vendors could enable system configurations which require that data is, to the extent possible within clinical practice, recorded in a structured way (in contrast to, e.g., free text fields). As the input data for process mining typically originates from several information systems, healthcare organizations are encouraged to take measures to ensure that data from different sources can be correlated correctly, e.g. by using a common patient identifier.

At the level of logging attitudes and mechanisms, healthcare organizations could start with raising awareness among healthcare professionals regarding

[2] An example is the free online course on process mining in healthcare, which is accessible at `https://www.futurelearn.com/courses/process-mining-healthcare`.

the importance of accurate and timely data registration. In this respect, the intended use cases for process mining within the organization can be highlighted. These same use cases will also highlight which specific data points need to be recorded more accurately. Besides awareness creation, measures could also be taken to facilitate data registration. A commonly heard complaint is that it takes too much time to systematically login to the terminals present within each room to record certain actions. When this holds, data registration can be facilitated by, e.g., equipping the computers in the department with a badge scanning system. This will lead to more accurate and timely data registrations, without placing additional burdens on healthcare professionals. While the aforementioned measures are beneficial to enable the systematic use of process mining, healthcare organizations will probably only take such measures when they are convinced of the added value of process mining. Once this is the case, healthcare organizations and professionals will become partners of the process mining community in the quest to improve the quality of recorded process execution data.

4.3. Integrate process mining functionalities in existing health information systems (HOV-3)

When the community aims to achieve a more widespread use of process mining in healthcare organizations, a close partnership with health information systems vendors is required. Process mining functionalities should be integrated in existing health information systems such that healthcare professionals can seamlessly use them, ideally moving towards the concept of process-aware information systems envisaged by Dumas et al. [94]. When process mining can only be applied using a standalone tool, data needs to be exported from the health information systems to import it in this tool (potentially requiring data restructuring). This imposes significant barriers to the use of process mining for clinical and non-clinical decision support.

While ensuring this seamless use is a responsibility for health information systems vendors, it is important to recognize that vendors are typically demand-driven when determining which features to add to their systems. Hence, healthcare organizations can motivate vendors to incorporate process mining functionalities by, e.g., stressing the need for data-driven process analysis functions or by incorporating it as a system requirement.

5. Conclusions

Despite the unique potential of process mining to retrieve data-driven insights in healthcare processes, its uptake in healthcare organizations is rather limited. Given this observation, this paper synthesized the outcomes of a two-day international brainstorm seminar on how to enhance the usability and understandability of process mining in healthcare. Based on the discussions of 18 experts, both researchers and healthcare practitioners, a set of key recommendations is specified. Ten recommendations target process mining researchers and the research community. While this was the predominant focus of the work, three additional recommendations are directed to healthcare organizations and health information systems vendors.

The recommendations to process mining researchers give rise to several research challenges. These include (i) setting up a benchmarking study to identify the most suitable process modeling language to visualize the output of control-flow discovery algorithms in healthcare, (ii) developing techniques to handle data quality issues in healthcare event logs, (iii) designing techniques which leverage different sources of process-related data, e.g. originating from IoT settings, which potentially originate from different healthcare organizations, (iv) creating process mining methods which approach healthcare processes from various perspectives and attribute central importance to the patient journey, and (v) guiding healthcare organizations in their process mining endeavors by developing a comprehensive methodology for process mining in healthcare. Besides these challenges, continued research efforts are required at the intersection of process mining and complementary techniques within areas such as predictive analytics, operations research, and machine learning.

Enhancing the usability and understandability of process mining in healthcare will require continued efforts by individual researchers and the research community. Some specific topics for future work have been highlighted in the previous paragraph. However, to strengthen the relevance of process mining in healthcare and to further increase its potential benefits, researchers are strongly encouraged to carefully consider the recommendations formulated in this paper in all their research projects. This also involves developing an attitude in which, for instance, real-world healthcare problems are a starting point for research, and healthcare specificities are taken into account when developing new process mining techniques. Besides the efforts of individual researchers, the research community will also play a key role. The community can, for instance, target the development of a standardized terminology for process mining in healthcare, or ensure that sufficient training opportunities for healthcare professionals are available. These united efforts will contribute to process mining reaching its full potential as a catalyst for evidence-based process improvement in healthcare.

References

[1] K. Kirchner, N. Herzberg, A. Rogge-Solti, M. Weske, Embedding conformance checking in a process intelligence system in hospital environments, Lecture Notes in Computer Science 7738 (2013) 126–139.

[2] R. S. Mans, W. M. P. van der Aalst, R. J. B. Vanwersch, Process mining in healthcare: evaluating and exploiting operational healthcare processes, Springer, Heidelberg, 2015.

[3] M. Dumas, M. La Rosa, J. Mendling, H. A. Reijers, Fundamentals of business process management, Springer, Heidelberg, 2013.

[4] E. Rojas, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Process mining in healthcare: a literature review, Journal of Biomedical Informatics 61 (2016) 224–236.

[5] R. Lenz, M. Reichert, IT support for healthcare processes–premises, challenges, perspectives, Data & Knowledge Engineering 61 (2007) 39–58.

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

[7] Z. Huang, W. Dong, L. Ji, L. Yin, H. Duan, On local anomaly detection and analysis for clinical pathways, Artificial Intelligence in Medicine 65 (2015) 167–177.

[8] C. Di Ciccio, A. Marrella, A. Russo, Knowledge-intensive processes: characteristics, requirements and analysis of contemporary approaches, Journal on Data Semantics 4 (2015) 29–57.

[9] S. Mertens, F. Gailly, G. Poels, Towards a decision-aware declarative process modeling language for knowledge-intensive processes, Expert Systems with Applications 87 (2017) 316–334.

[10] R. Lenz, M. Peleg, M. Reichert, Healthcare process support: achievements, challenges, current research, International Journal of Knowledge-Based Organizations 2 (2012).

[11] A. Najjar, D. Reinharz, C. Girouard, C. Gagné, A two-step approach for mining patient treatment pathways in administrative healthcare databases, Artificial Intelligence in Medicine 87 (2018) 34–48.

[12] W. M. P. van der Aalst, Process mining: data science in action, Springer, Heidelberg, 2016.

[13] A. Rule, M. F. Chiang, M. R. Hribar, Using electronic health record audit logs to study clinical activity: a systematic review of aims, measures, and methods, Journal of the American Medical Informatics Association 27 (2020) 480–490.

[14] Institute of Medicine, The Learning Healthcare System, Technical Report, 2007.

[15] A. Budrionis, J. G. Bellika, The learning healthcare system: where are we now? A systematic review, Journal of Biomedical Informatics 64 (2016) 87–92.

[16] C. P. Friedman, A. K. Wong, D. Blumenthal, Achieving a nationwide learning health system, Science Translational Medicine 2 (2010) 57cm29.

[17] R. S. Mans, W. M. P. van der Aalst, N. C. Russell, P. J. M. Bakker, A. J. Moleman, Process-aware information system development for the healthcare domain-consistency, reliability, and effectiveness, Lecture Notes in Business Information Processing 43 (2010) 635–646.

[18] D. Russler, Clinical event, in: L. Liu, M. T. Özsu (Eds.), Encyclopedia of Database Systems, Springer, New York, 2016.

[19] J. De Weerdt, F. Caron, J. Vanthienen, B. Baesens, Getting a grasp on clinical pathway data: an approach based on process mining, Lecture Notes in Artificial Intelligence 7769 (2013) 22–35.

[20] A. W. Kempa-Liehr, C. Y.-C. Lin, R. Britten, D. Armstrong, J. Wallace, D. Mordaunt, M. O'Sullivan, Healthcare pathway discovery and probabilistic machine learning, International Journal of Medical Informatics (2020) 104087.

[21] J. De Weerdt, M. De Backer, J. Vanthienen, B. Baesens, A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs, Information Systems 37 (2012) 654–676.

[22] A. Augusto, R. Conforti, M. Dumas, M. La Rosa, F. M. Maggi, A. Marrella, M. Mecella, A. Soo, Automated discovery of process models from event logs: review and benchmark, IEEE Transactions on Knowledge and Data Engineering 31 (2018) 686–705.

[23] Z. Huang, X. Lu, H. Duan, On mining clinical pathway patterns from medical behaviors, Artificial Intelligence in Medicine 56 (2012) 35–50.

[24] L. Chiudinelli, A. Dagliati, V. Tibollo, S. Albasini, N. Geifman, N. Peek, J. H. Holmes, F. Corsi, R. Bellazzi, L. Sacchi, Mining post-surgical care processes in breast cancer patients, Artificial Intelligence in Medicine (2020) 101855.

[25] W. M. P. van der Aalst, H. A. Reijers, M. Song, Discovering social networks from event logs, Computer Supported Cooperative Work 14 (2005) 549–593.

[26] A. Syamsiyah, B. F. van Dongen, W. M. P. van der Aalst, Discovering social networks instantly: moving process mining computations to the database and data entry time, Lecture Notes in Business Information Processing 287 (2017) 51–67.

[27] A. Pika, M. Leyer, M. T. Wynn, C. J. Fidge, A. H. T. Hofstede, W. M. V. D. Aalst, Mining resource profiles from event logs, ACM Transactions on Management Information Systems 8 (2017) 1–30.

[28] M. Song, W. M. P. van der Aalst, Towards comprehensive support for organizational mining, Decision Support Systems 46 (2008) 300–317.

[29] A. Burattin, A. Sperduti, M. Veluscek, Business models enhancement through discovery of roles, in: Proceedings of the 2013 IEEE Symposium on Computational Intelligence and Data Mining, IEEE, pp. 103–110.

[30] S. Suriadi, M. T. Wynn, J. Xu, W. M. van der Aalst, A. H. ter Hofstede, Discovering work prioritisation patterns from event logs, Decision Support Systems 100 (2017) 77–92.

[31] S. Mertens, F. Gailly, G. Poels, Discovering health-care processes using DeciClareMiner, Health Systems 7 (2018) 195–211.

[32] A. Bottrighi, F. Chesani, P. Mello, M. Montali, S. Montani, P. Terenziani, Conformance checking of executed clinical guidelines in presence of basic medical knowledge, Lecture Notes in Business Information Processing 100 (2011) 200–211.

[33] G. Geleijnse, H. Aklecha, M. Vroling, R. Verhoeven, F. N. van Erning, P. A. Vissers, J. C. Buijs, X. A. Verbeek, Using process mining to evaluate colon cancer guideline adherence with cancer registry data: a case study, in: Proceedings of the 2018 AMIA Annual Symposium.

[34] R. de la Fuente, R. Fuentes, J. Munoz-Gama, A. Riquelme, F. R. Altermatt, J. Pedemonte, M. Corvetto, M. Sepúlveda, Control-flow analysis of procedural skills competencies in medical training through process mining, Postgraduate Medical Journal 96 (2020) 250–256.

[35] J. Carmona, B. van Dongen, A. Solti, M. Weidlich, Conformance checking, Springer, Heidelberg, 2018.

[36] S. Dunzer, M. Stierle, M. Matzner, S. Baier, Conformance checking: a state-of-the-art literature review, in: Proceedings of the 11th International Conference on Subject-Oriented Business Process Management, pp. 1–10.

[37] W. M. P. van der Aalst, A. Adriansyah, B. van Dongen, Replaying history on process models for conformance checking and performance analysis, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2 (2012) 182–192.

[38] A. Rogge-Solti, W. M. P. van der Aalst, M. Weske, Discovering stochastic Petri nets with arbitrary delay distributions from event logs, Lecture Notes in Business Information Processing 171 (2014) 15–27.

[39] M. T. Wynn, E. Poppe, J. Xu, A. H. ter Hofstede, R. Brown, A. Pini, W. M. van der Aalst, Processprofiler3d: a visualisation framework for

log-based process performance comparison, Decision Support Systems 100 (2017) 93–108.

[40] A. Rozinat, W. M. P. van der Aalst, Decision mining in business processes, Technical Report BPM Center Report BPM-06-10, 2006.

[41] M. de Leoni, F. Mannhardt, Decision discovery in business processes, in: S. Sakr, A. Y. Zomaya (Eds.), Encyclopedia of Big Data Technologies, Springer, Heidelberg, 2019, pp. 614–625.

[42] H. Baek, M. Cho, S. Kim, H. Hwang, M. Song, S. Yoo, Analysis of length of hospital stay using electronic health records: A statistical and data mining approach, PloS One 13 (2018) e0195901.

[43] A. A. Boxwala, B. H. Rocha, S. Maviglia, V. Kashyap, S. Meltzer, J. Kim, R. Tsurikova, A. Wright, M. D. Paterno, A. Fairbanks, et al., A multi-layered framework for disseminating knowledge for computer-based decision support, Journal of the American Medical Informatics Association 18 (2011) i132–i139.

[44] M. Peleg, Computer-interpretable clinical guidelines: a methodological review, Journal of Biomedical Informatics 46 (2013) 744–763.

[45] N. Iglesias, J. M. Juarez, M. Campos, Comprehensive analysis of rule formalisms to represent clinical guidelines: selection criteria and case study on antibiotic clinical guidelines, Artificial Intelligence in Medicine 103 (2020) 101741.

[46] E. Batista, A. Solanas, Process mining in healthcare: a systematic review, in: Proceedings of the 9th International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1–6.

[47] T. G. Erdogan, A. Tarhan, Systematic mapping of process mining studies in healthcare, IEEE Access 6 (2018) 24543–24567.

[48] A. P. Kurniati, O. Johnson, D. Hogg, G. Hall, Process mining in oncology: a literature review, in: Proceedings of the 6th International Conference on Information Communication and Management (ICICM), pp. 291–297.

[49] G. P. Kusuma, M. Hall, C. P. Gale, O. A. Johnson, Process mining in cardiology: a literature review, Intern Journal of Bioscience, Biochemistry and Bioinformatics 8 (2018) 226–236.

[50] R. Williams, E. Rojas, N. Peek, O. A. Johnson, Process mining in primary care: a literature review, Studies in Health Technology and Informatics 247 (2018) 376–380.

[51] N. F. Farid, M. De Kamps, O. A. Johnson, Process mining in frail elderly care: A literature review, in: Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies, volume 5, pp. 332–339.

[52] 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 informatics research 19 (2013) 42–49.

[53] 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, Journal of Digital Imaging 29 (2016) 47–58.

[54] M. Vroling, Relating process aspects to survival for colon cancer patients in the Netherlands, Master's thesis, Eindhoven University of Technology, the Netherlands, 2018.

[55] R. Mans, H. Reijers, M. van Genuchten, D. Wismeijer, Mining processes in dentistry, in: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, pp. 379–388.

[56] E. Rojas, A. Cifuentes, A. Burattin, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Performance analysis of emergency room episodes through process mining, International Journal of Environmental Research and Public Health 16 (2019) 1274.

[57] S. Yoo, M. Cho, E. Kim, S. Kim, Y. Sim, D. Yoo, H. Hwang, M. Song, Assessment of hospital processes using a process mining technique: outpatient process analysis at a tertiary hospital, International Journal of Medical Informatics 88 (2016) 34–43.

[58] 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: Proceedings of eHealth Beyond the Horizon - Get IT There, pp. 573–578.

[59] S. Montani, G. Leonardi, S. Quaglini, A. Cavallini, G. Micieli, Improving structural medical process comparison by exploiting domain knowledge and mined information, Artificial Intelligence in Medicine 62 (2014) 33–45.

[60] A. Partington, M. Wynn, S. Suriadi, C. Ouyang, J. Karnon, Process mining for clinical processes: a comparative analysis of four australian hospitals, ACM Transactions on Management Information Systems 5 (2015) 1–18.

[61] S. Suriadi, R. Andrews, A. H. ter Hofstede, M. T. Wynn, Event log imperfection patterns for process mining: towards a systematic approach to cleaning event logs, Information Systems 64 (2017) 132–150.

[62] N. Martin, Using indoor location system data to enhance the quality of healthcare event logs: opportunities and challenges, Lecture Notes in Business Information Processing 342 (2018) 226–238.

[63] L. Reinkemeyer, Process mining in action: principles, use cases and outlook, Springer, Cham, 2020.

[64] R. Andrews, S. Suriadi, M. Wynn, A. H. ter Hofstede, S. Rothwell, Improving patient flows at St. Andrew's War Memorial Hospital's emergency department through process mining, in: J. vom Brocke, J. Mendling (Eds.), Business Process Management Cases, Springer, 2018, pp. 311–333.

[65] M. Engelhorn, Semantics and big data semantics methods for data processing and searching large amounts of data, in: P. Langkafel (Ed.), Big data in medical science and healthcare management, Walter de Gruyter, Berlin, 2016, pp. 177–196.

[66] World Health Organization, Classification of diseases, `https://www.who.int/classifications/icd/en/`, 2019. [Online; accessed 29-July-2019].

[67] World Health Organization, WHO releases new International Classification of Diseases (ICD 11), `https://www.who.int/news-`

room/detail/18-06-2018-who-releases-new-international-classification-of-diseases-(icd-11), 2018. [Online; accessed 29-July-2019].

[68] Regenstrief Institute, Logical Observation Identifiers Names and Codes, https://loinc.org/, 2020. [Online; accessed 15-February-2020].

[69] S. Welch, J. Augustine, C. A. Camargo Jr, C. Reese, Emergency department performance measures and benchmarking summit, Academic Emergency Medicine 13 (2006) 1074–1080.

[70] J. L. Wiler, S. Welch, J. Pines, J. Schuur, N. Jouriles, S. Stone-Griffith, Emergency department performance measures updates: proceedings of the 2014 emergency department benchmarking alliance consensus summit, Academic Emergency Medicine 22 (2015) 542–553.

[71] SNOMED International, 5-step briefing, http://www.snomed.org/snomed-ct/five-step-briefing, 2019. [Online; accessed 29-July-2019].

[72] L. Marco-Ruiz, K. Malm-Nicolaisen, R. Pedersen, A. Makhlysheva, P. A. Bakkevoll, Ontology-based terminologies for healthcare, Technical report, Norwegian Centre for E-health Research, Tromsø, Norway, 2017.

[73] OpenEHR, Open industry specifications, models and software for e-health, https://www.openehr.org, 2020. [Online; accessed 09-July-2020].

[74] G.-H. Ulriksen, R. Pedersen, G. Ellingsen, Infrastructuring in healthcare through the openehr architecture, Computer Supported Cooperative Work (CSCW) 26 (2017) 33–69.

[75] R. Pedersen, C. Granja, L. Marco-Ruiz, Implementation of openEHR in combination with clinical terminologies: experiences from Norway, International Journal on Advances in Life Sciences 9 (2017) 82–91.

[76] C. W. Günther, H. M. W. Verbeek, XES standard definition, Technical report, Eindhoven Unversity of Technology, Eindhoven, The Netherlands, 2014.

[77] XES Working Group and others, IEEE standard for eXtensible Event Stream (XES) for achieving interoperability in event logs and event streams, IEEE Std 1849-2016 (2016) 1–50.

[78] M. T. Wynn, W. Z. Low, A. H. ter Hofstede, W. Nauta, A framework for cost-aware process management: cost reporting and cost prediction, Journal of Universal Computer Science 20 (2014) 406–430.

[79] T. Baier, J. Mendling, M. Weske, Bridging abstraction layers in process mining, Information Systems 46 (2014) 123–139.

[80] T. J. Bright, A. Wong, R. Dhurjati, E. Bristow, L. Bastian, R. R. Coeytaux, G. Samsa, V. Hasselblad, J. W. Williams, M. D. Musty, et al., Effect of clinical decision-support systems: a systematic review, Annals of Internal Medicine 157 (2012) 29–43.

[81] D. Blum, S. X. Raj, R. Oberholzer, I. I. Riphagen, F. Strasser, S. Kaasa, E. IMPACT, et al., Computer-based clinical decision support systems and patient-reported outcomes: a systematic review, The Patient-Patient-Centered Outcomes Research 8 (2015) 397–409.

[82] P. Bennett, N. R. Hardiker, The use of computerized clinical decision support systems in emergency care: a substantive review of the literature, Journal of the American Medical Informatics Association 24 (2016) 655–668.

[83] P. Mazzocato, C. Savage, M. Brommels, H. Aronsson, J. Thor, Lean thinking in healthcare: a realist review of the literature, BMJ Quality & Safety 19 (2010) 376–382.

[84] A. D'Andreamatteo, L. Ianni, F. Lega, M. Sargiacomo, Lean in healthcare: a comprehensive review, Health Policy 119 (2015) 1197–1209.

[85] M. E. Porter, On competition, Harvard Business Press, Boston, 2008.

[86] C. Combi, M. Gozzi, B. Oliboni, J. M. Juarez, R. Marin, Temporal similarity measures for querying clinical workflows, Artificial Intelligence in Medicine 46 (2009) 37–54.

[87] L. Vanbrabant, K. Braekers, K. Ramaekers, I. Van Nieuwenhuyse, Simulation of emergency department operations:a comprehensive review of

KPIs and operational improvements, Computers & Industrial Engineering 131 (2019) 356–381.

[88] N. Martin, B. Depaire, A. Caris, The use of process mining in business process simulation model construction, Business & Information Systems Engineering 58 (2016) 73–87.

[89] B. Depaire, N. Martin, Data-driven process simulation, in: S. Sakr, A. Zomaya (Eds.), Encyclopedia of Big Data Technologies, Springer, Cham, 2018.

[90] M. Camargo, M. Dumas, O. González-Rojas, Automated discovery of business process simulation models from event logs, Decision Support Systems (2020) 113284.

[91] A. R. Hevner, S. T. March, J. Park, S. Ram, Design science in information systems research, MIS quarterly (2004) 75–105.

[92] P. Johannesson, E. Perjons, An introduction to design science, Springer, Heidelberg, 2014.

[93] J. R. Venable, Identifying and addressing stakeholder interests in design science research: an analysis using critical systems heuristics, IFIP Advances in Information and Communication Technology 301 (2009) 93–112.

[94] M. Dumas, W. M. P. van der Aalst, A. H. Ter Hofstede, Process-aware information systems: bridging people and software through process technology, John Wiley & Sons, Hoboken, 2005.

[95] W. M. P. van der Aalst, C. Stahl, Modeling business processes: a Petri Net-oriented approach, MIT Press, Cambridge, 2011.

[96] K. Figl, Comprehension of procedural visual business process models, Business & Information Systems Engineering 59 (2017) 41–67.

[97] G. De Giacomo, M. Dumas, F. M. Maggi, M. Montali, Declarative process modeling in BPMN, Lecture Notes in Computer Science 9097 (2015) 84–100.

[98] A. A. Andaloussi, A. Burattin, T. Slaats, E. Kindler, B. Weber, On the declarative paradigm in hybrid business process representations: a conceptual framework and a systematic literature study, Information Systems 91 (2020) 101505.

[99] S. J. Leemans, E. Poppe, M. T. Wynn, Directly follows-based process mining: Exploration & a case study, in: Proceedings of the 2019 International Conference on Process Mining, IEEE, pp. 25–32.

[100] J. Recker, Opportunities and constraints: the current struggle with BPMN, Business Process Management Journal 16 (2010) 181–201.

[101] M. Chinosi, A. Trombetta, BPMN: an introduction to the standard, Computer Standards & Interfaces 34 (2012) 124–134.

[102] M. Kocbek, G. Jošt, M. Heričko, G. Polančič, Business process model and notation: The current state of affairs, Computer Science and Information Systems 12 (2015) 509–539.

[103] M. Pesic, H. Schonenberg, W. M. P. van der Aalst, Declare: full support for loosely-structured processes, in: Proceedings of the 11th IEEE International Enterprise Distributed Object Computing Conference, pp. 287–287.

[104] T. Hildebrandt, R. R. Mukkamala, T. Slaats, Nested dynamic condition response graphs, Lecture Notes in Computer Science 7141 (2011) 343–350.

[105] T. T. Hildebrandt, R. R. Mukkamala, Declarative event-based workflow as distributed dynamic condition response graphs, Electronic Proceedings in Theoretical Computer Science 69 (2010) 59–73.

[106] J. De Smedt, J. De Weerdt, J. Vanthienen, Fusion miner: process discovery for mixed-paradigm models, Decision Support Systems 77 (2015) 123–136.

[107] Y. Shahar, S. Miksch, P. Johnson, The asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines, Artificial Intelligence in Medicine 14 (1998) 29–51.

[108] M. Peleg, A. A. Boxwala, O. Ogunyemi, Q. Zeng, S. Tu, R. Lacson, E. Bernstam, N. Ash, P. Mork, L. Ohno-Machado, et al., GLIF3: the evolution of a guideline representation format, in: Proceedings of the American Medical Informatics Association Symposium, pp. 645–649.

[109] D. R. Sutton, J. Fox, The syntax and semantics of the pro forma guideline modeling language, Journal of the American Medical Informatics Association 10 (2003) 433–443.

[110] M. Peleg, S. Tu, J. Bury, P. Ciccarese, J. Fox, R. A. Greenes, R. Hall, P. D. Johnson, N. Jones, A. Kumar, et al., Comparing computer-interpretable guideline models: a case-study approach, Journal of the American Medical Informatics Association 10 (2003) 52–68.

[111] R. P. J. C. Bose, W. M. P. van der Aalst, I. Žliobaitė, M. Pechenizkiy, Dealing with concept drifts in process mining, IEEE Transactions on Neural Networks and Learning Systems 25 (2013) 154–171.

[112] A. Maaradji, M. Dumas, M. La Rosa, A. Ostovar, Detecting sudden and gradual drifts in business processes from execution traces, IEEE Transactions on Knowledge and Data Engineering 29 (2017) 2140–2154.

[113] A. J. M. M. Weijters, J. T. S. Ribeiro, Flexible heuristics miner (FHM), in: Proceedings of the 2011 IEEE Symposium on Computational Intelligence and Data Mining, pp. 310–317.

[114] B. F. A. Hompes, J. C. A. M. Buijs, W. M. P. van der Aalst, P. M. Dixit, J. Buurman, Discovering deviating cases and process variants using trace clustering, in: Proceedings of the 27th Benelux Conference on Artificial Intelligence.

[115] F. Mannhardt, M. de Leoni, H. A. Reijers, W. M. P. van der Aalst, Data-driven process discovery-revealing conditional infrequent behavior from event logs, Lecture Notes in Computer Science 10253 (2017) 532–544.

[116] G. Ibanez-Sanchez, C. Fernandez-Llatas, A. Martinez-Millana, A. Celda, J. Mandingorra, L. Aparici-Tortajada, Z. Valero-Ramon, J. Munoz-Gama, M. Sepúlveda, E. Rojas, et al., Toward value-based healthcare through interactive process mining in emergency rooms: the

stroke case, International Journal of Environmental Research and Public Health 16 (2019) 1783.

[117] P. M. Dixit, H. Caballero, B. F. A. Hompes, J. C. A. M. Buijs, W. M. P. van der Aalst, Enabling interactive process analysis with process mining and visual analytics, in: Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies, pp. 573–584.

[118] N. Martin, A. Martinez-Millana, B. Valdivieso, C. Fernández-Llatas, Interactive data cleaning for process mining: a case study of an outpatient clinic's appointment system, Lecture Notes in Business Information Processing 362 (2019) 532–544.

[119] P. M. Dixit, J. C. A. M. Buijs, W. M. P. van der Aalst, B. F. A. Hompes, J. Buurman, Using domain knowledge to enhance process mining results, Lecture Notes in Business Information Processing 244 (2017) 76–104.

[120] P. M. Dixit, J. C. A. M. Buijs, W. M. P. van der Aalst, Prodigy: human-in-the-loop process discovery, in: Proceedings of the 12th International Conference on Research Challenges in Information Science, IEEE, pp. 1–12.

[121] G. Guyatt, J. Cairns, D. Churchill, D. Cook, B. Haynes, J. Hirsh, J. Irvine, M. Levine, M. Levine, J. Nishikawa, D. Sackett, P. Brill-Edwards, H. Gerstein, J. Gibson, R. Jaeschke, A. Kerigan, A. Neville, A. Panju, A. Detsky, M. Enkin, P. Frid, M. Gerrity, A. Laupacis, V. Lawrence, J. Menard, V. Moyer, C. Mulrow, P. Links, A. Oxman, J. Sinclair, P. Tugwell, Evidence-based medicine: a new approach to teaching the practice of medicine, JAMA 268 (1992) 2420–2425.

[122] B. Djulbegovic, G. H. Guyatt, Progress in evidence-based medicine: a quarter century on, The Lancet 390 (2017) 415–423.

[123] A. Levin, The cochrane collaboration, Annals of Internal Medicine 135 (2001) 309–312.

[124] R. Andrews, M. T. Wynn, K. Vallmuur, A. H. ter Hofstede, E. Bosley, M. Elcock, S. Rashford, Leveraging data quality to better prepare

for process mining: an approach illustrated through analysing road trauma pre-hospital retrieval and transport processes in queensland, International Journal of Environmental Research and Public Health 16 (2019) 1138.

[125] 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 (2015) 29821–29840.

[126] L. Vanbrabant, N. Martin, K. Ramaekers, K. Braekers, Quality of input data in emergency department simulations: framework and assessment techniques, Simulation Modelling Practice and Theory 91 (2019) 83–101.

[127] A. P. Kurniati, E. Rojas, D. Hogg, G. Hall, O. A. Johnson, The assessment of data quality issues for process mining in healthcare using medical information mart for intensive care iii, a freely available e-health record database, Health Informatics Journal (2018).

[128] P. M. Dixit, S. Suriadi, R. Andrews, M. T. Wynn, A. H. ter Hofstede, J. C. A. M. Buijs, W. M. P. van der Aalst, Detection and interactive repair of event ordering imperfection in process logs, Lecture Notes in Computer Science 10816 (2018) 274–290.

[129] A. Solti, Event log cleaning for business process analytics, in: S. Sakr, A. Zomaya (Eds.), Encyclopedia of Big Data Technologies, Springer, Heidelberg, 2018.

[130] A. Rogge-Solti, R. S. Mans, W. M. P. van der Aalst, M. Weske, Repairing event logs using timed process models, Lecture Notes in Computer Science 8186 (2013) 705–708.

[131] D. Bayomie, I. M. Helal, A. Awad, E. Ezat, A. ElBastawissi, Deducing case ids for unlabeled event logs, Lecture Notes in Business Information Processing 256 (2016) 242–254.

[132] C. Di Francescomarino, C. Ghidini, S. Tessaris, I. V. Sandoval, Completing workflow traces using action languages, Lecture Notes in Computer Science 9097 (2015) 314–330.

[133] F. Fox, V. R. Aggarwal, H. Whelton, O. Johnson, A data quality framework for process mining of electronic health record data, in: Proceedings of the 2018 IEEE International Conference on Healthcare Informatics, pp. 12–21.

[134] M. Najafabadipour, M. Zanin, A. Rodríguez-González, M. Torrente, B. N. García, J. L. C. Bermudez, M. Provencio, E. Menasalvas, Reconstructing the patient's natural history from electronic health records, Artificial Intelligence in Medicine (2020) 101860.

[135] E. V. Epure, P. Martín-Rodilla, C. Hug, R. Deneckère, C. Salinesi, Automatic process model discovery from textual methodologies, in: Proceedings of the 9th IEEE International Conference on Research Challenges in Information Science, IEEE, pp. 19–30.

[136] H. van der Aa, J. Carmona Vargas, H. Leopold, J. Mendling, L. Padró, Challenges and opportunities of applying natural language processing in business process management, in: Proceedings of the 27th International Conference on Computational Linguistics, Association for Computational Linguistics, pp. 2791–2801.

[137] E. V. Epure, D. Compagno, C. Salinesi, R. Deneckere, M. Bajec, S. Žitnik, Process models of interrelated speech intentions from online health-related conversations, Artificial Intelligence in Medicine 91 (2018) 23–38.

[138] S. Laumer, C. Maier, F. T. Gubler, Chatbot acceptance in healthcare: Explaining user adoption of conversational agents for disease diagnosis, in: Proceedings of the 27th European Conference on Information Systems.

[139] A. Piau, R. Crissey, D. Brechemier, L. Balardy, F. Nourhashemi, A smartphone chatbot application to optimize monitoring of older patients with cancer, International Journal of Medical Informatics 128 (2019) 18–23.

[140] L. Atzori, A. Iera, G. Morabito, The internet of things: a survey, Computer networks 54 (2010) 2787–2805.

[141] C. Janiesch, A. Koschmider, M. Mecella, B. Weber, A. Burattin, C. Di Ciccio, A. Gal, U. Kannengiesser, F. Mannhardt, J. Mendling, A. Oberweis, M. Reichert, S. Rinderle-Ma, W. Song, J. Su, V. Torres, M. Weidlich, M. Weske, L. Zhang, The internet-of-things meets business process management: mutual benefits and challenges, Technical report, arXiv:1709.03628, 2017.

[142] H. H. Nguyen, F. Mirza, M. A. Naeem, M. Nguyen, A review on IoT healthcare monitoring applications and a vision for transforming sensor data into real-time clinical feedback, in: Proceedings of the 2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design, IEEE, pp. 257–262.

[143] H. Ahmadi, G. Arji, L. Shahmoradi, R. Safdari, M. Nilashi, M. Alizadeh, The application of internet of things in healthcare: a systematic literature review and classification, Universal Access in the Information Society (2018) 1–33.

[144] T. Proesmans, C. Mortelmans, R. Van Haelst, F. Verbrugge, P. Vandervoort, B. Vaes, Mobile phone–based use of the photoplethysmography technique to detect atrial fibrillation in primary care: diagnostic accuracy study of the FibriCheck app, JMIR mHealth and uHealth 7 (2019) e12284.

[145] N. Goodwin, Understanding integrated care, International Journal of Integrated Care 16 (2016) 6.

[146] M. A. C. Bautista, M. Nurjono, Y. W. Lim, E. Dessers, H. J. M. Vrijhoef, Instruments measuring integrated care: a systematic review of measurement properties, The Milbank Quarterly 94 (2016) 862–917.

[147] C. C. Agbo, Q. H. Mahmoud, J. M. Eklund, Blockchain technology in healthcare: a systematic review, Healthcare 7 (2019) 56.

[148] P. Zhang, D. C. Schmidt, J. White, G. Lenz, Blockchain technology use cases in healthcare, in: P. Raj, G. C. Deka (Eds.), Advances in Computers, volume 111, Elsevier, 2018, pp. 1–41.

[149] J. Mendling, I. Weber, W. V. D. Aalst, J. V. Brocke, C. Cabanillas, F. Daniel, S. Debois, C. D. Ciccio, M. Dumas, S. Dustdar, et al.,

Blockchains for business process management - challenges and opportunities, ACM Transactions on Management Information Systems 9 (2018) 1–16.

[150] R. Mühlberger, S. Bachhofner, C. Di Ciccio, L. García-Bañuelos, O. López-Pintado, Extracting event logs for process mining from data stored on the blockchain, Lecture Notes in Business Information Processing 362 (2019) 690–703.

[151] M. A. Hailemichæl, K. Y. Yigzaw, J. G. Bellika, Emnet: a system for privacy-preserving statistical computing on distributed health data, in: Proceedings from The 13th Scandinavien Conference on Health Informatics, 115, pp. 33–40.

[152] F. Mannhardt, Multi-perspective process mining, Ph.D. thesis, Eindhoven University of Technology, 2018.

[153] T. M. Trebble, N. Hansi, T. Hydes, M. A. Smith, M. Baker, Process mapping the patient journey: an introduction, BMJ 341 (2010) c4078.

[154] M. Berghout, J. van Exel, L. Leensvaart, J. M. Cramm, Healthcare professionals' views on patient-centered care in hospitals, BMC Health Services Research 15 (2015).

[155] W. M. P. van der Aalst, A. Adriansyah, A. K. A. d. Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. v. d. Brand, R. Brandtjen, J. Buijs, A. Burattin, J. Carmona, M. Castellanos, J. Claes, J. Cook, N. Costantini, F. Curbera, E. Damiani, M. d. Leoni, P. Delias, B. F. v. Dongen, M. Dumas, S. Dustdar, D. Fahland, D. R. Ferreira, W. Gaaloul, F. van Geffen, S. Goel, C. Günther, A. Guzzo, P. Harmon, A. t. Hofstede, J. Hoogland, J. E. Ingvaldsen, K. Kato, R. Kuhn, A. Kumar, M. L. Rosa, F. Maggi, D. Malerba, R. S. Mans, A. Manuel, M. McCreesh, P. Mello, J. Mendling, M. Montali, H. R. Motahari-Nezhad, M. zur Muehlen, J. Munoz-Gama, L. Pontieri, J. Ribeiro, A. Rozinat, H. S. Pérez, R. S. Pérez, M. Sepúlveda, J. Sinur, P. Soffer, M. Song, A. Sperduti, G. Stilo, C. Stoel, K. Swenson, M. Talamo, W. Tan, C. Turner, J. Vanthienen, G. Varvaressos, E. Verbeek, M. Verdonk, R. Vigo, J. Wang, B. Weber, M. Weidlich, T. Weijters, L. Wen, M. Westergaard, M. Wynn, Process mining manifesto, Lecture Notes in Business Information Processing 99 (2012) 169–194.

[156] D. A. Cook, M. T. Teixeira, B. S. Heale, J. J. Cimino, G. Del Fiol, Context-sensitive decision support (infobuttons) in electronic health records: a systematic review, Journal of the American Medical Informatics Association 24 (2017) 460–468.

[157] A. K. Teng, A. B. Wilcox, A review of predictive analytics solutions for sepsis patients, Applied Clinical Informatics 11 (2020) 387–398.

[158] A. Perotte, R. Ranganath, J. S. Hirsch, D. Blei, N. Elhadad, Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis, Journal of the American Medical Informatics Association 22 (2015) 872–880.

[159] D. Delen, B. Davazdahemami, E. Eryarsoy, L. Tomak, A. Valluru, Using predictive analytics to identify drug-resistant epilepsy patients, Health Informatics Journal 26 (2020) 449–460.

[160] Y. Chen, T. A. Lasko, Q. Mei, J. C. Denny, H. Xu, A study of active learning methods for named entity recognition in clinical text, Journal of Biomedical Informatics 58 (2015) 11–18.

[161] M. Bozkaya, J. Gabriels, J. M. van der Werf, Process diagnostics: a method based on process mining, in: Proceedings of the 2009 International Conference on Information, Process, and Knowledge Management, IEEE, pp. 22–27.

[162] M. L. van Eck, X. Lu, S. J. J. Leemans, W. M. P. van der Aalst, PM^2: a process mining project methodology, Lecture Notes in Computer Science 9097 (2015) 297–313.

[163] C. Fernández-Llatas, T. Meneu, V. Traver, J.-M. Benedi, Applying evidence-based medicine in telehealth: an interactive pattern recognition approximation, International journal of environmental research and public health 10 (2013) 5671–5682.

[164] E. Rojas, M. Sepúlveda, J. Munoz-Gama, D. Capurro, V. Traver, C. Fernandez-Llatas, Question-driven methodology for analyzing emergency room processes using process mining, Applied Sciences 7 (2017) 302.

[165] F. Caron, J. Vanthienen, K. Vanhaecht, E. Van Limbergen, J. De Weerdt, B. Baesens, A process mining-based investigation of adverse events in care processes, Health Information Management Journal 43 (2014) 16–25.

[166] R. A. Q. Neira, B. F. A. Hompes, J. G.-J. de Vries, B. F. Mazza, S. L. S. de Almeida, E. Stretton, J. C. Buijs, S. Hamacher, Analysis and optimization of a sepsis clinical pathway using process mining, Lecture Notes in Business Information Processing 362 (2019) 459–470.

[167] T. K. Colicchio, J. J. Cimino, G. Del Fiol, Unintended consequences of nationwide electronic health record adoption: challenges and opportunities in the post-meaningful use era, Journal of Medical Internet Research 21 (2019) e13313.