Journal of Biomedical Informatics 127 (2022) 103994

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

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

Original Research

# Process mining for healthcare: Characteristics and challenges

Jorge Munoz-Gama [a,*], Niels Martin [b,c,*], Carlos Fernandez-Llatas [d,g,*], Owen A. Johnson [e,*], Marcos Sepúlveda [a,*], Emmanuel Helm [f,*], Victor Galvez-Yanjari [a,*], Eric Rojas [a], Antonio Martinez-Millana [d], Davide Aloini [k], Ilaria Angela Amantea [l,q,r], Robert Andrews [ab], Michael Arias [z], Iris Beerepoot [o], Elisabetta Benevento [k], Andrea Burattin [ai], Daniel Capurro [j], Josep Carmona [s], Marco Comuzzi [w], Benjamin Dalmas [aj,ak], Rene de la Fuente [a], Chiara Di Francescomarino [h], Claudio Di Ciccio [i], Roberto Gatta [ad,ae], Chiara Ghidini [h], Fernanda Gonzalez-Lopez [a], Gema Ibanez-Sanchez [d], Hilda B. Klasky [p], Angelina Prima Kurniati [al], Xixi Lu [o], Felix Mannhardt [m], Ronny Mans [af], Mar Marcos [v], Renata Medeiros de Carvalho [m], Marco Pegoraro [x], Simon K. Poon [ag], Luise Pufahl [u], Hajo A. Reijers [m,o], Simon Remy [y], Stefanie Rinderle-Ma [ah], Lucia Sacchi [t], Fernando Seoane [g,am,an], Minseok Song [aa], Alessandro Stefanini [k], Emilio Sulis [l], Arthur H. M. ter Hofstede [ab], Pieter J. Toussaint [ac], Vicente Traver [d], Zoe Valero-Ramon [d], Inge van de Weerd [o], Wil M.P. van der Aalst [x], Rob Vanwersch [m], Mathias Weske [y], Moe Thandar Wynn [ab], Francesca Zerbato [n]

[a] *Pontificia Universidad Católica de Chile, Chile*
[b] *Hasselt University, Belgium*
[c] *Research Foundation Flanders (FWO), Belgium*
[d] *Universitat Politècnica de València, Spain*
[e] *University of Leeds, United Kingdom*
[f] *University of Applied Sciences Upper Austria, Austria*
[g] *Karolinska Institutet, Sweden*
[h] *Fondazione Bruno Kessler, Italy*
[i] *Sapienza University of Rome, Italy*
[j] *University of Melbourne, Australia*
[k] *University of Pisa, Italy*
[l] *University of Turin, Italy*
[m] *Eindhoven University of Technology, The Netherlands*
[n] *University of St. Gallen, Switzerland*
[o] *Utrecht University, The Netherlands*
[p] *Oak Ridge National Laboratory, United States*
[q] *University of Bologna, Italy*
[r] *University of Luxembourg, Luxembourg*
[s] *Universitat Politècnica de Catalunya, Spain*
[t] *University of Pavia, Italy*
[u] *Technische Universitaet Berlin, Germany*
[v] *Universitat Jaume I, Spain*
[w] *Ulsan National Institute of Science and Technology (UNIST), Republic of Korea*
[x] *RWTH Aachen University, Germany*
[y] *University of Potsdam, Germany*
[z] *Universidad de Costa Rica, Costa Rica*
[aa] *Pohang University of Science and Technology, Republic of Korea*
[ab] *Queensland University of Technology, Australia*
[ac] *Norwegian University of Science and Technology, Norway*
[ad] *Universitá degli Studi di Brescia, Italy*

* Corresponding authors at: Hasselt University, Belgium (N. Martin), Karolinska Institutet, Sweden (C. Fernandez-Llatas), Pontificia Universidad Católica de Chile (J. Munoz-Gama, M. Sepúlveda, V. Galvez-Yanjari), Research Foundation Flanders (FWO), Belgium (N. Martin), Universitat Politècnica de València, Spain (C. Fernandez-Llatas), University of Applied Sciences Upper Austria, Austria (E. Helm) and University of Leeds (O. Johnson).

https://doi.org/10.1016/j.jbi.2022.103994
Received 1 September 2021; Received in revised form 16 November 2021; Accepted 10 January 2022
Available online 29 January 2022

[ae] Lausanne University Hospital (CHUV), Switzerland
[af] Philips Research, the Netherlands
[ag] The University of Sydney, Australia
[ah] Technical University of Munich, Germany
[ai] Technical University of Denmark, Denmark
[aj] Mines Saint-Etienne, France
[ak] Université Clermont Auvergne, France
[al] Telkom University, Indonesia
[am] Karolinska University Hospital, Sweden
[an] University of Borås, Sweden

ARTICLE INFO

*Keywords:*
Process mining
Healthcare

ABSTRACT

Process mining techniques can be used to analyse business processes using the data logged during their execution. These techniques are leveraged in a wide range of domains, including healthcare, where it focuses mainly on the analysis of diagnostic, treatment, and organisational processes. Despite the huge amount of data generated in hospitals by staff and machinery involved in healthcare processes, there is no evidence of a systematic uptake of process mining beyond targeted case studies in a research context. When developing and using process mining in healthcare, distinguishing characteristics of healthcare processes such as their variability and patient-centred focus require targeted attention. Against this background, the Process-Oriented Data Science in Healthcare Alliance has been established to propagate the research and application of techniques targeting the data-driven improvement of healthcare processes. This paper, an initiative of the alliance, presents the distinguishing characteristics of the healthcare domain that need to be considered to successfully use process mining, as well as open challenges that need to be addressed by the community in the future.

## 1. Introduction

Innovations make healthcare better, affordable and efficient. Developments such as new technologies and business models help to move healthcare forward [1]. In addition, healthcare systems are confronted across the world with unprecedented challenges, including the permanent and rapid adaptation of clinical processes based on the emerging scientific evidence [2] and the provision of high-quality care with limited resources [3–5]. Within this context, healthcare organisations, such as hospitals, are aware of the need to manage and improve both their clinical processes (e.g., care pathways describing the treatment of a particular medical condition over the time) and their organisational/administrative processes (e.g., billing processes) [6,7].

Process execution data are a valuable information source to support the management and improvement of healthcare processes. Typically, healthcare organisations make intensive use of Health Information Systems (HISs), such as a hospital information system. During the execution of a process, several entries in the HISs are recorded (e.g., when a patient was registered or was subject to a clinical examination by a physician). The entries in the databases of these HISs can be leveraged to generate an event log describing the sequence of activities that were performed, when they were executed, by whom and for whom (e.g., for which specific patient) [8]. As the event log reflects how a process has been executed in reality, it can support clinicians, healthcare organisation's managers, and other decision-makers with a wide range of process-related questions in the medical domain.

To answer process-related questions, *process mining* techniques can be of great value. Process mining is a set of techniques used in many domains, including healthcare, to retrieve valuable insights from an event log [7–9]. A multitude of process mining techniques have been developed in industry and academia [8,10], which enable healthcare stakeholders to identify the actual order of activities in a process [11], to determine the conformance between an existing (e.g., normative) model and reality [12], and to provide insights into the involvement of resources in a process [8,13].

Compared to alternative approaches, such as process mapping exercises with staff members [14], process mining takes data about the real-life behaviour of a process as a starting point. In this way, process mining can support healthcare institutions in achieving each of the quadruple aim for healthcare improvement [15]: (i) improving population health, e.g., by supporting the analysis and improvement of care pathways; (ii) improving patient experience, e.g., by highlighting how a process can be streamlined from the patient's perspective; (iii) reducing costs, e.g., by making bottlenecks explicit; and (iv) improving the work-life balance of healthcare professionals, e.g., by enabling the analysis of resource involvement and requirements in a healthcare process [15,16]. Besides supporting the data-driven management and improvement of healthcare processes, process mining also has the potential to support the resilience of the healthcare system by providing detailed insights into how processes are being executed within a particular context [2].

Despite the vast potential of process mining [17], the distinguishing characteristics of the healthcare domain require targeted attention when developing and using process mining techniques. For instance: healthcare processes are typically characterised by high levels of variation [18], and strongly depend on complex decisions made by knowledge workers, such as physicians, which often have a high degree of autonomy [19,20]. To date, there is no evidence of a systematic uptake of process mining beyond targeted case studies in a research context [7]. To induce a widespread and systematic adoption of process mining in the healthcare domain, targeted methods and techniques that explicitly take the specific domain characteristics into account are required.

Our paper aims to convey a unified perspective on the distinguishing characteristics and associated key challenges of the *Process Mining for Healthcare (PM4H)* domain. The distinguishing characteristics add to the complexity of using process mining within a healthcare context. The associated challenges need to be tackled by the community, including researchers and practitioners, in order to structurally establish process mining in healthcare as a powerful instrument to support evidence-based process analysis and improvement. Tackling the highlighted challenges requires efforts at the level of fundamental research and translational research. In that sense, this paper may inspire and encourage researchers to contribute to this exciting field.

This paper is an initiative of the *Process-Oriented Data Science for Healthcare Alliance*[1], which is affiliated with the *IEEE Task Force on Process Mining*[2]. The alliance aims to promote research and knowledge

[1] https://www.pods4h.com/alliance
[2] https://www.tf-pm.org

sharing to encourage targeted innovation in process mining in healthcare, as well as its effective application in real-life settings. With this paper, the alliance expresses its firm ambition to strengthen the domain's identity by making both its distinguishing characteristics and key challenges explicit. The distinguishing characteristics and challenges have emerged from extensive consultation with a wide range of experts on the topic, for instance during panel discussions at the PODS4H workshop[3] at the International Conference on Business Process Management and the International Conference on Process Mining.

The remainder of this paper is structured as follows. Section 2 introduces the basic concepts of process mining in healthcare. Section 3 outlines ten distinguishing characteristics which are especially relevant for process mining in healthcare, and Section 4 presents ten associated challenges that have to be considered by the community. Table 1 presents an overview of the distinguishing characteristics and challenges in the PM4H domain. Finally, the paper ends with some concluding remarks in Section 5.

## 2. Process Mining in Healthcare: Basic Concepts and Applications

This section introduces the key concepts of process mining (Section 2.1). Moreover, some applications of process mining in healthcare are presented (Section 2.2).

### 2.1. Process Mining

Process mining is a family of techniques focused on gaining valuable insights from data that processes generate while being executed. It works as a bridge between process science (which includes areas such as business process management and operations research) and data science (which includes areas such as data mining and predictive analytics), resulting in methods to analyse processes through data [8]. Process mining is domain-agnostic, i.e, process mining techniques can be applied in any industry where processes are present and data representing them is available. Healthcare, the focus of this paper, is one domain where the use of process mining is growing [10,21].

Fig. 1 positions process mining in healthcare in a wider context. Process mining centres around processes, which can be represented using a process model such as a model representing the sequence of steps in a process and the different paths a process can take [22]. The order of activities in a process can be visualised in various ways, e.g., using flowcharts [23] or Business Process Modeling Notation (BPMN) [24]. Fig. 2 exemplifies the use of BPMN to model the trajectory of sepsis patients. Note that the model is purposefully simplified for illustrative purposes.

Nowadays, many healthcare processes are – at least partially – supported by Health Information Systems (HISs). HISs that commonly support healthcare processes are Electronic Health Records (EHR) systems provided by vendors such as Epic[4], Cerner[5], MEDITECH[6], Allscripts[7], athenahealth[8], IBM[9], McKesson[10], and Siemens[11]. Such systems record data about the execution of processes in a healthcare organisation [8]. This process execution data can be leveraged to create an event log.

Event logs containing process execution data are the primary input for process mining algorithms [8]. An event log is composed of cases representing different process instances, e.g. the execution of a treatment process for a specific patient. Each case is composed of a sequence of events, where an event could refer to the completion of a particular activity in the treatment process. As illustrated in Table 2, an event log typically records the following information for each event: (a) an identifier of each case ('Case id'), (b) the activities that each case included ('Activity'), and (c) a reference to when each activity was executed ('Timestamp'). Besides this information, an event log can also contain information regarding the type of event ('Transaction type'), the resource associated to an event ('Resource'), as well as other attributes regarding the activity or case. Consequently, the key components of an event log are:

- *Case id*: a process instance contained in the data, which could be a patient being treated in the hospital in a clinical process or an invoice

**Table 1**
Distinguishing characteristics and challenges of process mining in healthcare.

| | Distinguishing Characteristics | | Challenges |
|---|---|---|---|
| D1: | **Exhibit Substantial Variability**<br>Variability is the norm due to intrinsic differences in patients. | C1: | **Design Dedicated/Tailored Methodologies and Frameworks**<br>PM4H methodologies and frameworks are essential to provide answers to relevant healthcare questions. |
| D2: | **Value the Infrequent Behaviour**<br>Infrequent behaviour contains key information for healthcare process understanding. | C2: | **Discover Beyond Discovery**<br>Conformance Checking and Enhancement types can provide new knowledge to the PM4H domain. |
| D3: | **Use Guidelines and Protocols**<br>Plenty of protocols and guidelines are present in healthcare. | C3: | **Mind the Concept Drift**<br>Healthcare processes change over the time. |
| D4: | **Break the Glass**<br>Context-related issues and unpredictable situations shape healthcare processes. | C4: | **Deal with Reality**<br>If it does not work with real data, it does not work. |
| D5: | **Consider Data at Multiple Abstraction Levels**<br>High and low level data are registered while healthcare processes are executed. | C5: | **Do It Yourself (DIY)**<br>Developments to foster the use of process mining by healthcare professionals are required. |
| D6: | **Involve a Multidisciplinary Team**<br>All the healthcare professionals implicated have key process knowledge. | C6: | **Pay Attention to Data Quality**<br>Addressing quality issues make results trustable. |
| D7: | **Focus on the Patient**<br>A patient-centred view is the core in healthcare processes. | C7: | **Take Care of Privacy and Security**<br>Methods to preserve patients' privacy and security when conducting PM4H analyses are needed. |
| D8: | **Think about White-box Approaches**<br>Understandability is key when analysing healthcare processes. | C8: | **Look at the Process through the Patient's Eyes**<br>New knowledge can emerge looking the process from the patient' perspective. |
| D9: | **Generate Sensitive and Low-Quality Data**<br>Healthcare data are sensitive, which are either automatically or manually recorded. | C9: | **Complement HISs with the Process Perspective**<br>How to incorporate the process perspective in the HIS is a challenge to be addressed. |
| D10: | **Handle Rapid Evolutions and New Paradigms**<br>Healthcare is constantly evolving due to advances in clinical knowledge, technology, and novel paradigms. | C10: | **Evolve in Symbiosis with the Developments in the Healthcare Domain**<br>PM4H methods should provide ongoing support to the constantly changing healthcare domain. |

[3] https://www.pods4h.com
[4] https://www.epic.com
[5] https://www.cerner.com
[6] https://ehr.meditech.com
[7] https://ca.allscripts.com
[8] https://www.athenahealth.com
[9] https://www.ibm.com/products/explorys-ehr-data-analysis-tools
[10] https://www.mckesson.com
[11] https://www.corporate.siemens-healthineers.com

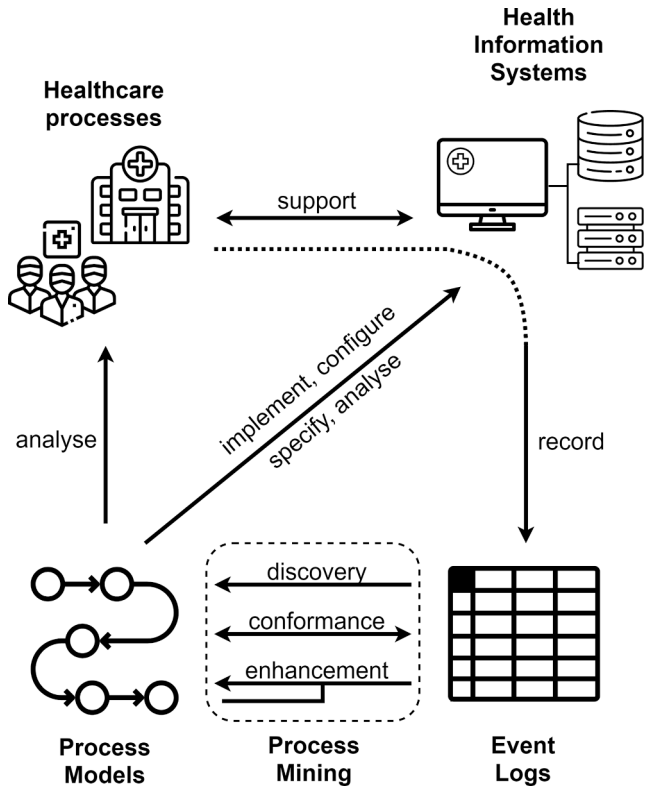


**Fig. 1.** Positioning of process mining in healthcare, based on [8].

in an organisational process. In Table 2, each case represents a sepsis patient. Cases can be recognized by their unique identifier, e.g. patient 253 and 255.

- *Activity*: a step of the process, which could be 'Check vital signs' in a clinical process or 'Send payment reminder' in an organisational process. In Table 2, the activities include 'ER Registration', representing the action of registering the patient in the emergency ward, and 'CRP', representing the registration of c-reactive protein measurement in the system.
- *Timestamp*: time at which the event took place. In Table 2, the first event took place on April 14th, 2021 at 11:33:50.
- *Transaction type*: the state when the event was recorded. In Table 2, all the events were recorded once they were completed ('complete' state). Other possible states are 'start', 'resume' or 'schedule', among others.
- *Resource*: refers to the resource associated to the event. This could refer, for instance, to a physician, a healthcare professional or medical device. In Table 2, the resource information tells us, for example, that 'Nurse 1' has registered the triage information of patient 253, while 'Physician 02' recorded the admission of patient 253 to the normal care ward.
- Other attributes: additional case or event attributes may also be recorded in an event log, such as the hospital unit where the patient has received care, the patient's vital signs, etc.

Event log creation in healthcare environments faces considerable challenges due to heterogeneous data sources (e.g., mobile health data) and distributed healthcare providers [27]. Moreover, it is recognised that some healthcare processes are not directly supported by HISs, i.e. *unplugged* processes [28]. These type of processes require novel methods to obtain an event log, such as the video tagging method to create process data from surgical procedure recordings [29]. Medical vocabularies and ontologies are also important assets that can be helpful when creating event logs in healthcare environments [27].

Using the event log, various process mining types can be performed in order to generate valuable process-related insights. Three prominent types of process mining are [8]:

- *Discovery:* these algorithms are useful to obtain process models reflecting process behavior from an event log [8]. Many algorithms focus on discovering the order of activities in the process (a.k.a. control-flow, trajectories, activity paths, and care pathways) from the data. Besides the large number of control-flow discovery algorithms, other discovery algorithms help to gain knowledge about how resources work throughout the process, focusing on role discovery [30,31], social networks [32,33], and task prioritisation patterns [34]. Examples of use cases include the discovery of models of gynecological oncology processes showing relations between the organisational units involved and therapeutics pathways that patients received [35,36], or the discovery of collaboration patterns discovered between the physician, nurse and dietitian involved in diabetes treatment [37]. Fig. 3 illustrates the discovered trajectories of sepsis patients in a hospital.
- *Conformance checking:* these algorithms require a process model, either obtained by means of a discovery algorithm or previously designed, and aim to compare the behaviour in the event log with the behaviour in that process model [12,36]. Hence, conformance checking algorithms help to evaluate whether the process is being executed as described in the process model, as well as to detect deviations between the observed behaviour in the event log and the process model. For instance: Huang et al. [9] use conformance checking to detect deviations from the unstable angina pathway. The analysis highlighted the presence of unexpected, early, delayed and absent activities [9].
- *Enhancement:* these algorithms help to enrich and extend an existing process model using process data [8]. An enhancement type is model repair, which allows the modification of a process model based on event logs [38]. Another type is model extension, where information is added to enrich a process model with information such as time and roles. Some examples are the repair of declarative process models representing clinical practice guidelines in an urology department [39], and the model extension of a process model with the duration of surgical procedure stages based on Robot-Assisted Partial Nephrectomy (RAPN) data [40].

To facilitate the application of process mining in practice, various software tools have been developed, including commercial software such as Disco[12], Celonis[13] and Apromore[14]. There are also open-source alternatives, such as ProM[15], and libraries for programming languages, such as pm4py[16] for Python and bupaR[17] for R.

### 2.2. Applications of Process Mining in Healthcare

Over the last years, process mining has been used for various use cases in healthcare, which has been demonstrated in various literature reviews on this topic [10,21,41]. While Section 2.1 present various process mining examples, this subsection highlights some of the most relevant healthcare questions that process mining can answer.

Process variant exploration for clinical pathway analysis is a frequently recurring use case for PM4H [21]. These techniques study the difference between several variants of the execution of the same clinical pathway, which could be approached from a control-flow or performance perspective [42]. Clinical pathways are highly flexible as all

[12] https://www.fluxicon.com/disco
[13] https://www.celonis.com
[14] https://www.apromore.org
[15] https://www.promtools.org
[16] https://pm4py.org
[17] https://www.bupar.net

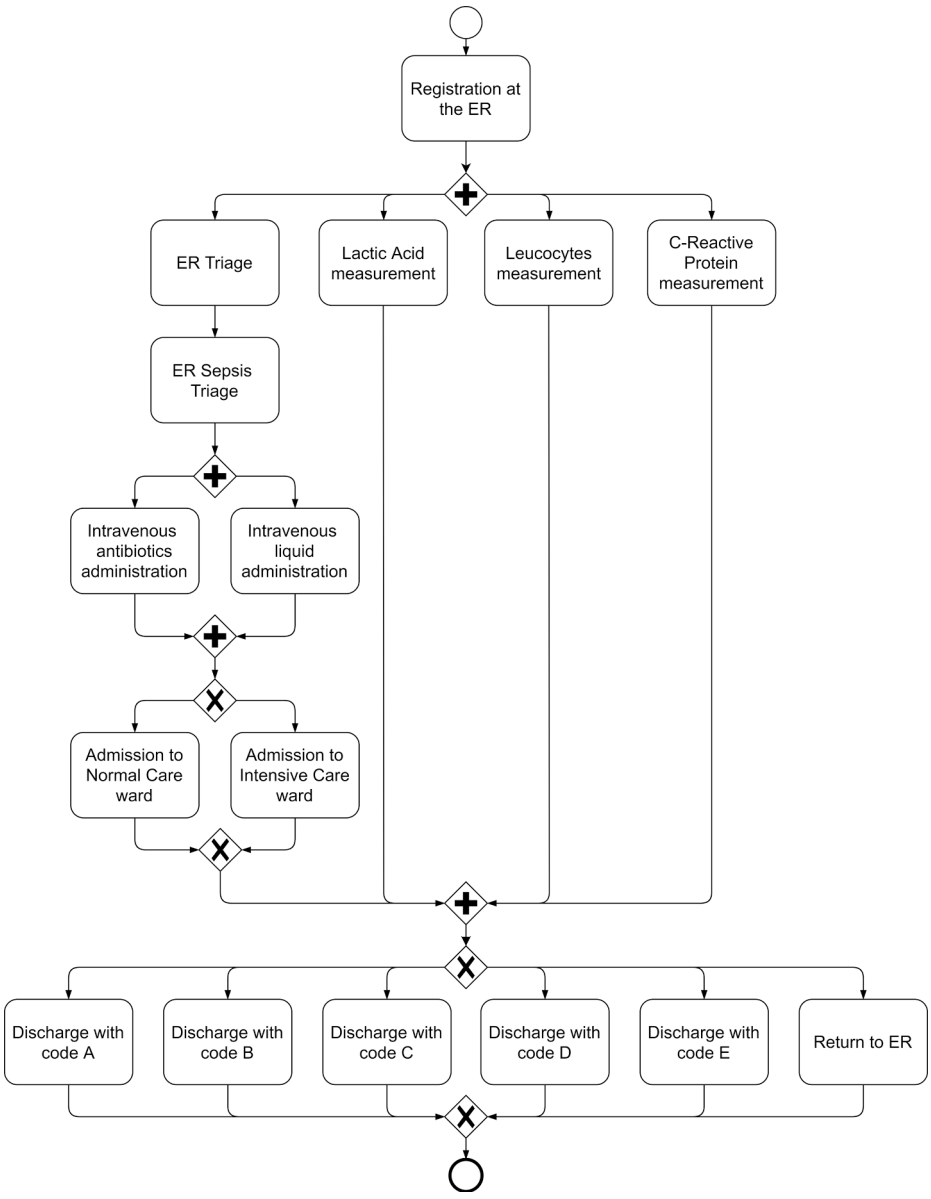


**Fig. 2.** Process model representing the sepsis patients trajectory, based on Mannhardt [25]. The model was created using BPMN as a process modelling language. The start event (○) indicates the start point of the process, and the end event (circle at the end of the model) indicates the end point of the process. The gateways represent alternative paths: the parallel gateway (✚) means that all the paths should be followed, and the exclusive gateway (✕) means that only one path .can be followed.

patients in need of the same treatment come with different co-morbidities and complications, involve complex decision-making due to its knowledge-intensive nature, are performed by a network of specialists, and continuously evolve due to innovations and unforeseen situations [43]. Identifying differences between groups of pathway executions through process variant analysis helps to decide whether process improvement is needed, and if so, which changes can make the process more efficient [44]. An example of this analysis is Caron et al. [45], where data of 1143 gynecologic oncology patients were analysed in two subsets: one with patients receiving radiotherapy and the other with patients receiving chemotherapy. There are various challenges related to process variant analysis, including the comparison of process variants from the resource perspective, the verification of guideline compliance, the discovery of how adverse events are faced, the analysis of process variants across the whole patient journey (including prevention, pre-hospital care, hospital treatment and rehabilitation), and the identification of useful variants for process improvement [42–46].

Process mining also has applications for disease trajectory modeling, which refers to models characterising the progress of a disease over time and compare the disease evolution depending on patient attributes such as the age, co-morbidities and received treatments received of a patient [47,48]. An example is the study conducted by De Oliveira et al. [49], where data of 76.523 sepsis patients was used to uncover the most common diagnostics that patients received prior and after the diagnosis to understand better how to identify and manage sepsis. Using process mining, they developed a bow-tie visualisation, which allowed them to discover that pneumonia and gastrointestinal disorders commonly occurred before sepsis, while septicaemia, pneumonia and urinary tract infections occurred after sepsis. Challenges in disease trajectory modeling using process mining are the development of models which are easy to understand (i.e. clear visualisations of the trajectories), and the comparison of these models with clinical guidelines using conformance checking [48].

The aforementioned use cases only aim to illustrate the opportunities that process mining offers to healthcare. There are various other highly relevant questions for which process mining can generate relevant

**Table 2**
Sepsis patients *event log* example, based on [26].

| Case id | Activity | Timestamp | Transaction type | Resource | … |
|---|---|---|---|---|---|
| … | … | … | … | … | |
| 253 | ER Triage | 04–13-2021 11:33:50 | complete | Nurse 1 | … |
| 255 | Release A | 04–13-2021 11:35:05 | complete | Physician 02 | … |
| 259 | Lactic Acid | 04–13-2021 11:38:55 | complete | Nurse 4 | … |
| 254 | Leucocytes | 04–13-2021 11:41:23 | complete | Nurse 5 | … |
| 256 | Lactic Acid | 04–13-2021 11:52:35 | complete | Nurse 4 | … |
| 257 | ER Triage | 04–13-2021 11:53:16 | complete | Nurse 7 | … |
| 258 | ER Registration | 04–13-2021 11:54:47 | complete | Nurse 8 | … |
| 253 | Admission NC | 04–13-2021 11:55:26 | complete | Physician 02 | … |
| 259 | Admission IC | 04–13-2021 11:58:30 | complete | Physician 03 | … |
| 260 | CRP | 04–13-2021 12:01:12 | complete | Nurse 07 | … |
| 261 | Release B | 04–13-2021 12:02:00 | complete | Physician 03 | … |
| 253 | IV Liquid | 04–13-2021 12:05:33 | complete | Nurse 2 | … |
| … | … | … | … | … | … |

insights. These questions include:

- *How does the process flow of patients with a particular medical complication differ from other patients?*
  Every patient is unique, which implies that patients with the same illness respond differently to the same treatment due to co-morbidities and other contextual factors [50]. Variations in the patient trajectory can be discovered with process mining algorithms, which help to characterise groups of similar patients (in terms of medical history, laboratory tests, etc.), allowing healthcare professionals to gain profound insights in the treatment trajectory of various patient types.
- *To which extent is the care pathway for a particular medical condition followed in practice?*
  With the rise of evidence-based medicine, protocols and clinical guidelines are developed to provide clarity in the necessary steps when diagnosing and treating a medical condition [51]. However, it is difficult to determine the implementation and effectiveness of clinical protocols and guidelines in reality, i.e. whether they are followed in practice. Process mining allows practitioners and researchers to perform this type of analysis, which can help to understand major deviations from clinical guidelines, as well as to identify areas for improvement in clinical guidelines and protocols.
- *Where are the bottlenecks in a healthcare process?*
  Time is often an important variable in healthcare. Process mining makes it possible to analyse the time perspective of processes through indicators such as waiting times and activity duration, which together help to detect bottlenecks in a process. Having this information on healthcare processes, such as those in an emergency department, can drive decision-making to, for example, improve the availability of boxes and reduce waiting times [52].
- *How do multiple clinical experts interact in a care process?*
  Collaboration between clinicians and other healthcare staff is daily practice in healthcare [53]. Hence, when analysing a care process, various healthcare professionals are likely to be involved when treating a disease. Process mining provides tools to analyse collaboration patterns among healthcare professionals within a process, e.g. by identifying handovers of work [54].

These questions illustrate that PM4H can support healthcare professionals in answering a wide variety of process-related questions. Against this background, the next section will outline distinguishing characteristics of healthcare processes. Afterwards, key challenges for the PM4H community are discussed.

## 3. Distinguishing Characteristics

This section outlines ten distinguishing characteristics of healthcare processes, which have implications for PM4H. While we do not claim that these characteristics are exclusive to the healthcare domain, we consider them as highly relevant for the use of process mining in a healthcare context. The distinguishing characteristics are discussed separately in the remainder of this section, but, in practice, they are also interconnected, adding to the complexity required to take them into consideration. Moreover, the distinguishing characteristics give rise to specific challenges, which need to be taken into account when developing process mining techniques.

### 3.1. D1: Exhibit Substantial Variability

Healthcare processes are complex, in part because they tend to exhibit significant variability [4,18,55]. Several factors contribute to this intrinsic presence of variability in healthcare processes. These factors include the vast diversity of activities that can typically be executed, the fact that several subprocesses can be executed simultaneously (e.g., in case of polytrauma), and the influence of differences in the personal preferences/characteristics of patients, clinicians and other healthcare professionals (e.g., impacting choices made in the treatment process) [18,56]. The combination of such factors tends to make almost all cases (e.g., a patient in a clinical process) different. For instance: given the patient's pathologies and co-morbidities, a different set of activities might need to be executed in comparison with the standard pathway. Moreover, patients can respond very differently to particular treatments, which affects the order or type of activities that follow. It should also not be forgotten that the patient is the ultimate decision maker, who may accept or decline a particular treatment according to beliefs, fears or perceptions regarding quality of life.

When an event log of a highly variable healthcare process is used to discover a control-flow model, control-flow discovery algorithms are likely to generate an unstructured model, often referred to as a *spaghetti model* [8]. Classic process mining techniques are not well prepared to deal with unstructured processes and, as a consequence, generate process models which are extremely challenging to interpret. A common approach to deal with this issue is to remove or reduce the variability in the event log by means of abstraction techniques such as filtering or aggregation, e.g., using trace clustering techniques [57] and semantic aggregation of activities [58,59]. However, this approach generates process models that only cover a small part of the problem at hand. Such approaches might not be sufficient for many real-world healthcare applications because they only provide a partial view of the process and may hide valuable infrequent behaviour. Hence, PM4H researchers should be aware of the variability issue when providing solutions, tools and frameworks to understand and deal with this variability.

### 3.2. D2: Value the Infrequent Behaviour

While infrequent behaviour could be considered as noise in a general scenario, it can be a source of valuable knowledge in the healthcare domain. Healthcare is known for being especially prone to *workarounds*, i.e., intentional deviations from prescribed practices [60]. Therefore, infrequent behaviour typically needs to be considered in PM4H. For example, nurses must check the vital signs of a patient before a consultation with a physician, and should immediately register the scores in the HIS. However, an analysis of the process might show that nurses keep track of the scores on a notepad and insert all the

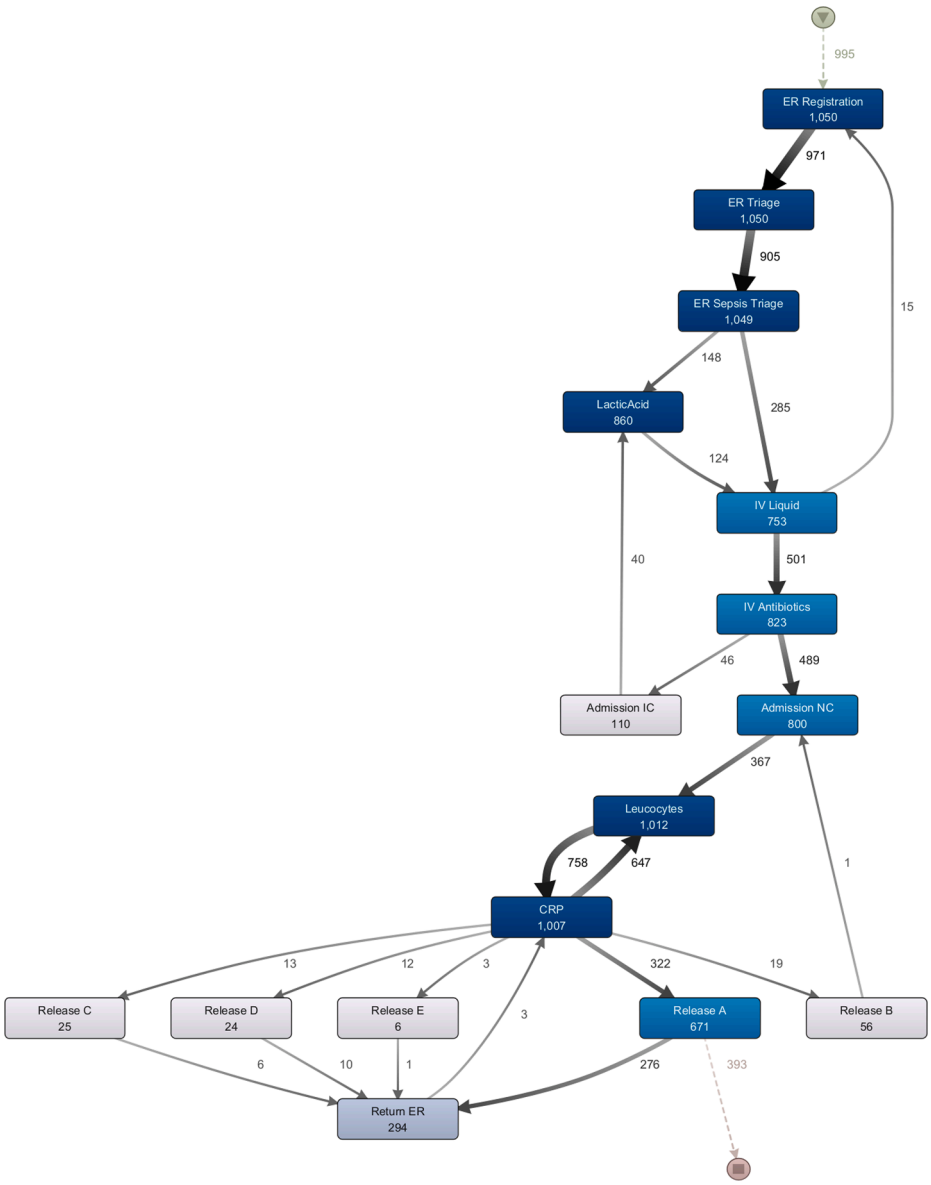


**Fig. 3.** Process map discovered using the Disco software and the sepsis patients data [26].

information later in their shifts [61,62]. Such workarounds provide insights in common inefficiencies and obstacles that healthcare actors face in their daily work. These insights provide a basis for a thorough analysis, enabling healthcare organisations to improve their processes [63]. In the aforementioned example, a shortage of computers and significant time pressure before the consultation with a physician could explain workarounds, providing valuable input for improvement actions [61]. Workarounds can also highlight that different paths through a treatment process lead to the same outcome, providing information about relevant treatment variations to treat a particular disease [64].

It follows that PM4H researchers and practitioners must go beyond simply filtering out infrequent behaviour from the event log. In contrast, they should try to understand why infrequent behaviour is observed and what this could mean. They should be aware that focusing on models of the typical execution of the process can result in blind spots, causing them to miss important opportunities for process improvement.

### 3.3. D3: Use Guidelines and Protocols

Over the past decades, evidence-based medicine has emerged as a process in which researchers and practitioners in the field of medicine continuously search for unbiased clinically-relevant information [51]. The evidence-based medicine paradigm has been accompanied by an increasing presence of *clinical practice guidelines and protocols* [65,66]. Clinical practice guidelines are systematically structured mechanisms which assist practitioners in determining the appropriate healthcare procedures for specific clinical circumstances [67] and, hence, constitute reference processes. The abundant presence of guidelines and protocols provides rich opportunities for process mining in the healthcare domain, compared to other domains where reference processes are absent [68]. These opportunities are twofold. Firstly, guidelines and protocols can facilitate the application of process mining. For instance, the available guidelines can be seen as prescriptive models to which the actual execution of the process can be compared using conformance checking algorithms [69] and similarity-based techniques [70–73]. In this way, PM4H practitioners can use the structured information provided by guidelines to select relevant events during event log generation [27], and to formally define the activities – with the associated activity labels – that will be considered by process discovery algorithms when generating the models [74,75]. Secondly, process mining can provide

the evidence required to improve guidelines and protocols by assessing their effectiveness and efficiency. Similarly, this generated evidence can be used to customise existing global standardisation efforts to the specific characteristics of local contexts in an affordable and effective way [76].

In summary, PM4H researchers and practitioners should be aware of the existing guidelines and protocols in their specific setting, such that they can be maximally leveraged during the preparation and analysis of an event log.

### 3.4. D4: Break the Glass

*'D3: Use Guidelines and Protocols'* highlights that standardised guidelines and protocols aim to establish high-level structure for healthcare processes. However, following the metaphor of a fire alarm, physicians and healthcare professionals might need to *break the glass* and deviate from the protocols that are in place. Such situations can occur both at the level of an individual patient and at the level of the system. At the patient level, alternative courses of action might need to be taken due to previously unknown co-morbidities, unexpected complications, patient preferences, or because certain (combinations of) co-morbidities are not covered by the existing guidelines [77–79]. At the system level, emergency situations might follow from a sudden surge in the number of incoming patients or an unforeseen reduction in staff availability [78]. Thus, healthcare processes deal with unplanned situations due to unforeseen emergencies or the optimization of the limited resources available.

When physicians and healthcare professionals react to unforeseen or emergency situations, this is, of course, also reflected in the data. This behaviour gives rise to unexpected patterns in the data, which stresses the importance of considering contextual information when conducting process mining analyses in healthcare [80]. Contextual information, such as patient characteristics or the state of the system at a particular point in time, can be essential to explain patterns observed in the data. This understanding can enable a PM4H researcher to systematically analyse the desirability of deviations from protocols, taking into account rich contextual information. Hence, considering the context can highlight the need to fine-tune protocols or provide input for developing future policies of a healthcare institution.

In sum, PM4H researchers and practitioners need to be aware of the existence of *break the glass* situations. Moreover, building upon the available contextual information, they can try to identify and understand such situations.

### 3.5. D5: Consider Data at Multiple Abstraction Levels

It is a common misconception that PM4H only uses medical treatment data and, hence, is a synonym for medical treatment analysis. Medical treatment analysis is, without doubt, an important use case for PM4H [10], with a multitude of examples being reported in medical domains such as oncology [81], cardiology [82], primary care [83], and frail elderly care [84]. However, a wide spectrum of different types of data is available in the healthcare domain, both related to clinical processes as well as to organisational/administrative processes [6,7,85].

A key distinction can be made between high-level data and low-level data. Low-level data are very fine-grained data, which is recorded by medical equipment or sensors at healthcare facilities. Typically, some form of aggregation is required to retrieve meaningful patterns from large volumes of fine-grained data [86]. For instance, event data recorded by medical equipment, such as an Allura Xper X-ray system, can be used to construct realistic test profiles for fault diagnosis to identify new problems before they actually emerge [87]. Another source of low-level data are technologies such as wearable devices and Real-Time Localisation Systems. The deployment of these technologies by healthcare organisations has experienced a remarkable increase in recent years, and their data has been successfully used as input for process mining to analyse patient pathways [88], and to gain insights in personalised chronic disease management [89]. In surgical procedures, even when surgical robots have not been used, alternative sources of low-level data – such as video recordings – can be used to identify activities, in order to analyse a particular procedure [28,29].

High-level data, which are more coarse, typically allow obtaining more meaningful patterns during the analysis without the need for aggregation. Data that is typically recorded in a hospital information system or administrative data often have a more high-level character. Due to this high-level character, such data might be suitable for cross-organisational comparisons, e.g., to share best practices amongst healthcare organisations [90]. Knowledge sharing regarding processes can be especially relevant in situations such as the COVID-19 crisis [2]. Consequently, cross-organisational comparisons can be of interest for the PM4H discipline.

Given the multitude of high-level and low-level data sources available in a typical healthcare organisation, a PM4H researcher and practitioner is likely to be confronted with different types of data. Such data include logistics, billing, accounting, staff, and payroll information, and can be used to answer important questions in topics such as shift management, bed management, patient admission, transfer and diversion policies [91]. Moreover, as different stakeholders have different information needs, a PM4H research question under analysis might require the combination of very distinct data sources.

### 3.6. D6: Involve a Multidisciplinary Team

Healthcare processes are increasingly having a multidisciplinary character [37]. Similarly, data science is, in its nature, multidisciplinary [92], and process mining is no exception [8]. Any type of process analysis could possibly include techniques from other computer science disciplines, such as artificial intelligence, machine learning, and computer vision [93].

Besides the potential involvement of expertise from the aforementioned computer science fields, PM4H researchers and practitioners should always recognise the multidisciplinary nature of healthcare processes, necessitating the involvement of extensive expertise from the healthcare domain. In order to ensure the relevance of PM4H techniques, the involvement of physicians, nurses and other healthcare professionals from all relevant departments is essential. From a practical angle, a multidisciplinary team should be composed, which is closely involved in all stages of the process mining effort: data collection, data analysis, data interpretation, the communication of its results and its translation to practical actions.

When working in this multidisciplinary context, PM4H researchers and practitioners should take the following considerations into account. Firstly, the team should consider other clinicians, besides physicians. As an example, many insightful PM4H studies in the literature demonstrate a deep involvement of nurses in the generation of their findings [61]. Secondly, PM4H projects must use the appropriate medical language, terminology, codes and customs to ensure mutual understanding in a multidisciplinary context. In this regard, a broad range of standards and ontologies on medical concepts, such as medications, procedures, and diagnoses, already exist. Examples of standardised terminologies are well established clinical ontologies such as SNOMED CT and ICD-10 [94].

### 3.7. D7: Focus on the Patient

Many actors are involved in a typical healthcare process. These include patients, the patients' relatives, physicians, nurses, other healthcare professionals and supporting staff. As mentioned in *'D6: Involve a Multidisciplinary Team'*, different staff categories should be involved in a process mining project team. However, it needs to be recognised that the patient is at the heart of all healthcare processes. Hence, PM4H researchers should always make sure to look at a process

from the patient's perspective, even when they might not be explicitly represented in the project team. This point implies that the patient perspective should receive explicit attention when developing methods, tools, and frameworks. In this way, PM4H can support healthcare organisations into providing patient-centred care, a key indicator for care quality [95,96].

When considering the journey of a patient with a particular medical condition, it should be recognised that the patient typically crosses the boundaries of several healthcare organisations. Besides the hospital, a patient might also visit other professionals such as a general practitioner and a paramedic. PM4H researchers need to take this cross-organisational character of the patient journey into account as data will also be spread over various organisations [7]. Even when considering a process within a single healthcare organisation, a patient-centered approach requires a specific mindset as process mining analyses typically focus on gaining insights into processes at the system level. In contrast, clinicians tend to focus on care for individual patients, which can have implications on the performance of the system as a whole. For instance: when a physician wants to completely finish a procedure for one patient before assessing new patients, this might lead to increasing waiting times at the system level. This example highlights that, next to the system perspective, PM4H research should also be aware of the individual patient when studying processes. This analysis involves both the clinical perspective as well as the patient experience perspective.

### 3.8. D8: Think about White-box Approaches

Physicians face difficult situations and decisions in their daily routine. Despite their extensive training, decisions on the required diagnostics and treatments for a patient will rely on a risk–benefit assessment by a physician, which will be context-dependent. The unprecedented advances in artificial intelligence and machine learning have fostered new decision support systems delivering accurate information to support physicians when making clinical decisions. However, one of the biggest challenges is that physicians are reluctant to trust and adopt recommendations by a system that they do not fully understand, referring to it as a *black box* [97].

To enable the data-driven improvement of healthcare processes, there is a need for *white box* approaches that enable healthcare actors to understand the origins of particular observations. Process mining techniques are perceived as *white boxes* [98], since their final goal is not to provide a categorical answer, but to provide healthcare actors with techniques to get a better understanding of what is going on in their processes. The strength of process mining lies in its focus on *understandability*. Therefore, PM4H researchers and practitioners should be aware of the critical importance of understandability. In that respect, they should be prepared to use novel visualisation techniques and interactive models that advance the field in the direction of providing healthcare professionals with the insights they need in an understandable way [99]. Moreover, there are decisions that cannot always be made effectively based on few parameters and constructs (from simple switch-case patterns to complex DMN features). Those decision points can benefit from more understandable formalisms, such as Computer-Interpretable Clinical Guidelines (CIGs) [78] allows the enactment evidence-based recommendations intended to optimize patient care [100]. Finally, PM4H researchers and practitioners need to recognise that approaches that *recommend and explain* various courses of action are more likely to be accepted than counterparts that 'enforce' a particular one [101].

### 3.9. D9: Generate Sensitive and Low-Quality Data

When performing process mining in a healthcare context, in particular in a clinical setting, data sensitivity needs to be taken into account. The data at hand might include information such as a patient's current medical condition, co-morbidities, treatments, etc. As a consequence, healthcare-related event logs must be handled carefully. Healthcare data are well-known for being sensitive because of its confidentiality, and its usage, storage, and transfer is strictly regulated by institutions, countries, and even international treaties [102]. Therefore, as responsible citizens, PM4H practitioners must consider ethics in general and data privacy, either when conducting a primary (e.g., to improve patients outcomes) or secondary use (e.g., to improve health services outcomes) of the data [103].

Next to data privacy, poor data quality is also an important issue within the healthcare domain. In many countries, healthcare processes are still paper-based to a certain extent, which presents challenges to attach precise timestamps to activities that have been conducted. Even though integrated HISs are becoming more pervasive, this does not guarantee high-quality data [104]. As shown in several existing works, healthcare processes tend to be characterised by poor quality data [62,105,106]. A variety of data quality issues can be identified, including missing events, imprecise timestamps, and incorrect timestamps [5,62]. The presence of data quality problems can be attributed, at least partly, to the fact that event recording often still requires a manual action from clinicians or administrative staff. When, for instance, an activity is not recorded in the HISs at the time that it was performed, the timestamp of the associated event will not correspond with the activity execution time. Moreover, clinicians tend to develop their own habits in terms of system registrations, e.g. based on their personal interpretation of the situation and the registration options provided by the system. The latter can give rise to different registration patterns across clinicians involved in the same process.

From the previous, it follows that PM4H is intrinsically connected to the need for explicit attention to data privacy and data quality. Hence, PM4H researchers and practitioners need to take both of these concepts carefully into account when they have the ambition to support the management and improvement of real-life healthcare processes.

### 3.10. D10: Handle Rapid Evolutions and New Paradigms

Healthcare is a domain that exhibits rapid and continuous evolutions, which also has implications for healthcare processes. A prime example is the change in healthcare processes due to the accumulation of knowledge from clinical research [7,51,65,66]. This principle of evidence-based medicine is based on the critical appraisal of various types of scientific studies such as randomized controlled trials and cohort studies. High-quality research findings are integrated in clinical practice, which can induce changes in diagnostic or treatment processes [107,108]. In a similar vein, rapid evolutions in technology also impact processes in healthcare. For instance: mobile health solutions present various opportunities to reshape healthcare processes. This relates, amongst others, to remote monitoring, which enables healthcare professionals to follow-up specific clinical parameters remotely [109].

Healthcare does not only evolve due to new clinical knowledge or technological advances, but also due to the emergence of novel paradigms. For instance, the patient-centered care paradigm implies that care should carefully consider the needs, values and preferences of each individual patient [110]. Within this context, patient-reported measures are increasingly gaining attention. In this respect, a distinction is made between patient-reported outcome measures (PROMs), focusing on health aspects such as symptoms and treatment side effects, and patient-reported experience measures (PREMs), centering around experiences when receiving care [111]. Both PROMs and PREMs constitute highly valuable output to evaluate and redesign healthcare processes.

In order to provide valuable contributions to the healthcare domain, the PM4H community should be able to handle rapid evolutions and new paradigms. This requires a permanent awareness of new trends and innovations, as well as a careful assessment of their impact on healthcare processes and, hence, PM4H. Advances in the healthcare domain might give rise to new information needs, on which PM4H researchers and

practitioners should anticipate.

## 4. Challenges

The previous section outlined the distinguishing characteristics of healthcare processes. These characteristics give rise to challenges that need to be studied by the PM4H community – both researchers and practitioners – to structurally embed process mining in the healthcare domain as an instrument to support evidence-based process analysis and continuous improvement. At the research level, the outlined challenges will require fundamental and translational research, where the latter is needed to support the actual uptake of fundamental research in healthcare practice. The challenges are meant to guide both aspiring and established PM4H researchers in their search for relevant research endeavors. Similar to the distinguishing characteristics, we do not claim that the challenges are exclusive to the healthcare domain. However, we consider them particularly relevant to move PM4H forward in the future.

The remainder of this section will outline the key challenges for PM4H. Each subsection's title will also link the challenge to one or more of the distinguishing characteristics by using the labels introduced in Section 3.

### 4.1. C1: Design Dedicated/Tailored Methodologies and Frameworks (D1 – D9)

Given the distinguishing characteristics outlined in Section 3, the domain needs novel PM4H *methodologies* and *frameworks*, which guide researchers and practitioners through the various phases of a PM4H analysis. Such an analysis ranges from the identification of the research problem, through the composition of the event log (taking into account considerations regarding privacy and security), to the actual analysis and interpretation of the results [7]. New methodologies and frameworks should remain flexible, such that they can be adapted to the specific characteristics of a country, institution, department, process, clinician, or even the patient involved.

General process mining methodologies, such as L*[8], PM$^2$ [112] and Aguirre et al. [113], were key factors in the rising popularity of process mining, since they opened the door for researchers of other disciplines to apply process mining to a wide range of contexts [114]. In an analogous way, healthcare-specific methodologies for PM4H will allow healthcare actors to incorporate process mining into their analysis [115–117]. Moreover, such frameworks may allow PM4H techniques to be reused in different contexts, providing the means for a fair comparison among techniques across different scenarios [118]. Frameworks for cross-organisational [90] and cross-national [75] studies are also an interesting research line to advance the PM4H state-of-the-art, as well as the generation of methodologies to conduct meta-analyses based on outcomes and metrics obtained with PM4H techniques [119].

### 4.2. C2: Discover Beyond Discovery (D3, D5)

As highlighted in Section 2, process mining techniques are commonly classified into three types: *discovery*, *conformance*, and *enhancement* [8]. However, the evolution of these three types of techniques over time has not been the same [120]. Initially, most of the process mining techniques were discovery techniques, making it the dominant research stream in the early days of process mining. In later years, conformance checking has gradually started gaining prominence, and more recently, the number of novel enhancement techniques are starting to rise significantly. Recently, there has also been a surge of novel techniques, such as event and trace abstraction techniques [86], simulation [121], and predictive process monitoring [122].

PM4H manifests a similar evolution, where discovery is the most dominant use case in published PM4H literature [10,81–84,41]. Although works that focus on discovery are highly relevant, novel conformance checking and enhancement techniques, tailored towards healthcare, are needed. Some seminal works have been proposed that explore conformance for healthcare [123–125]. However, they are still in their infancy, making significant potential for future growth, especially because conformance checking approaches can be used to evaluate the various guidelines and protocols that exist in healthcare (*'D3: Use Guidelines and Protocols'*). Similarly, some enhancement techniques have been proposed for healthcare, but there is still a lot to explore [126], especially regarding highly relevant process perspectives for healthcare such as the time perspective [52,127] and the resource perspective [30]. These two perspectives appear to be particularly promising since they can support the long term resource planning and service (re-) design of healthcare organisations [128]. Process mining can also help to build advanced simulation and forecasting models for the operations of healthcare departments (e.g. the emergency department) or for diagnostic-therapeutic pathways (e.g. patients suffering from breast cancer) [115,129].

### 4.3. C3: Mind the Concept Drift (D1, D3, D4)

Clinical practice guidelines and protocols tend to change over time, e.g., due to advances in clinical research [7,51,65,66]. Even without changes to guidelines and protocols, the execution of healthcare processes could change, for instance, due to seasonal factors, to cope with the influx of patients in a hospital during winter months, or because an alternative way of working is introduced. Gaining insights into such dynamic processes is challenging and can benefit from techniques to detect changes in the process and determine their effects.

To this end, research on concept drift, which refers to the phenomenon whereby processes change over time, is essential [130–132]. The evolution of the COVID-19 treatment process over time is an obvious illustration of this phenomenon [133]. Moreover, distinguishing characteristics such as the high variability of healthcare processes also make it challenging to identify and study change patterns. While these aspects show the need for further dedicated research efforts, concept drift should systematically be taken into account in PM4H.

### 4.4. C4: Deal with Reality (D2, D5, D9)

*'In theory there is no difference between theory and practice, while in practice there is'*. This proverb seems to have particular significance for PM4H. Whenever a novel PM4H technique is being developed, thorough testing and evaluation is a key element [118,134]. Similar to any general process mining technique, synthetic data can be used during the development of a PM4H technique in order to evaluate its performance in a controlled environment. However, PM4H research has the goal to generate research with societal impact and, hence, research should focus on novel approaches able to handle real-life healthcare data. Moreover, the techniques should be able to handle large amounts of data and cope with input data containing significant variability. Carefully considering real-life healthcare data is crucial as it can significantly influence the premises and design of the PM4H domain. When a researcher does not have access to real-life data due to the absence of partnerships with healthcare institutions, publicly available healthcare datasets can be used instead[18]. In this respect, community efforts can also be made in the direction of providing and maintaining extensively documented, publicly accessible datasets to PM4H researchers and practitioners.

### 4.5. C5: Do It Yourself (DIY) (D6, D8)

As outlined in *'D6: Involve a Multidisciplinary Team'*, PM4H is a multidisciplinary domain that requires the involvement of physicians,

[18] Examples of public healthcare logs are available at https://www.pods4h.com/pods4h

nurses, and other healthcare professionals. To support a widespread use of process mining, healthcare professionals should be able to perform their analyses with little or no assistance of process mining experts. This direct involvement of healthcare actors has several implications that need to be taken into account when PM4H researchers propose new techniques. Firstly, during technique development, healthcare actors should be targeted as end users. Therefore, techniques should not require extensive expertise in terms of process mining in order to be used correctly. Secondly, the output of techniques should be understandable for healthcare actors. For instance, in case of control-flow discovery, the process modelling notation used to visualise the output should be appropriate for non-process mining experts from the healthcare domain [7]. To that end, simple visualisations, such as Directly-Follows Graphs (DFGs) [135] and Business Process modelling and Notation (BPMN) diagrams [136,137], may be effective. However, when using notations without formal semantics, it is important to consider their well-known limitations, such as the risk for misleading statistics and ambiguous visualizations [138]. Finally, all tools or software developed in PM4H research should be user-friendly, with specific attention to terminology, human–computer interaction, and visualisations.

### 4.6. C6: Pay Attention to Data Quality (D9)

The *'garbage in - garbage out'* principle also holds for PM4H, implying that the quality of all analyses ultimately depends upon the quality of the data used as input. As mentioned in *'D9: Generate Sensitive and Low-Quality Data'*, real-life data from HISs tends to suffer from data quality issues, which is troublesome for the use of these data for process mining purposes. The widespread presence of data quality issues stresses the need for the development of techniques to systematically assess and improve the quality of healthcare event logs. Recent approaches that have been developed, for example [62,139–142], demonstrate the relevance of this research topic. A recent overview on event log quality in PM4H is presented in Martin [106]. While existing approaches typically focus on data quality in the context of an existing event log, it should be recognised that data quality issues are also related to data management and extraction. This is because stakeholders typically have difficulties extracting this data [143].

Given the direct impact of data quality issues on the outcomes of process mining and the high prevalence of such issues in a healthcare context, it is a challenge to raise awareness about the topic within the healthcare sector [104]. Key players, such as hospital managers and physicians, should consider the impact of data quality issues on potential process mining analyses. Besides improving awareness, techniques to support the (interactive) identification of data quality issues, as well as their rectification (if possible) are valuable [144]. Moreover, analysis outcomes should also contain a reflection on the quality of the underlying data [7]. When interpreting the results, expressing the uncertainty degree of process mining outcomes can reflect the required level of caution on the part of healthcare actors. This aspect is especially important when process mining insights will be used within the context of clinical decision-making.

### 4.7. C7: Take Care of Privacy and Security (D9)

PM4H, especially when used in a clinical context, builds upon sensitive data regarding patients. Adequately safeguarding data privacy and security is of key importance for PM4H, which include responsible data science aspects such as fairness (avoiding prejudice) and confidentiality (not revealing sensitive information) [145]. In the broader process mining community, the importance of data privacy and security has already been recognised in the *Process Mining Manifesto* [146]. However, actual research contributions focusing on privacy-preserving techniques targeted to process-related data only started to appear recently [147,148]. It still remains difficult to balance safeguarding privacy on the one hand and preserving the value of the event log for process mining purposes on the other hand [102].

While anonymisation and other privacy-preserving techniques can support healthcare organisations when creating an event log suitable for data exchange (e.g., between a hospital and research partners), PM4H could also explore alternative *modus operandi* to enable the collaboration between researchers and healthcare organisations limiting the need for data exchange. Establishing methods to efficiently set up such collaborations can move PM4H forward, as it reduces the risk of data breaches. This point is especially relevant in times where data privacy and security concerns tend to make healthcare organisations more reluctant to share data for research purposes.

### 4.8. C8: Look at the Process through the Patient's Eyes (D7)

Within the context of a process mining project, information needs can differ depending on the involved stakeholder. As mentioned earlier, potential stakeholders in the healthcare domain include hospital managers, physicians, nurses, patients, and their relatives [7]. Currently, many process mining initiatives in healthcare target hospital management, heads of department or clinicians and often aim to provide a high-level overview of process behaviour. However, as reflected by particularity *'D7: Focus on the Patient'*, all processes directly or indirectly focus on the patient.

Against this background, there is a need for studying healthcare processes through the patient's eyes: what is the patient's journey while being diagnosed or treated for a particular (potentially chronic) condition? Studying a process from the patient's perspective can help physicians to consider the full patient journey when making decisions at a particular point in time. Even when adopting the patient's perspective does not directly result in options to improve the clinical or administrative outcomes, it might bring ways to the uncover on how to enhance the patient's experience [149,150]. Looking at the process from the patient's eyes also should consider the combination of data from various departments, and even different healthcare institutions [7]. In this way, process mining analyses will move from a single department/institution perspective to a cross department/institution perspective. This point of view constitutes a complementary perspective to existing process mining research in healthcare.

### 4.9. C9: Complement HISs with the Process Perspective (D2,D4,D5)

As important as focusing on the data and their uses for process mining, PM4H research must also focus on the source of the data, the Health Information Systems (HISs), which can include from more traditional hospital information systems [151,152] to more integrated information systems [153]. Such systems could benefit from being complemented with the process perspective, giving their processes a more predominant role, and interacting with process awareness systems to support both the operations of the organisation and the processes involved [154]. This combination will open the door for process mining to many of the benefits of systems with a strong process perspective –such as Process-Aware Information Systems (PAISs) [155]–, but will also require research and study of the inherent problems and future directions of both HISs [156] and their process awareness combination [157]. This study must include the understanding of how HISs must integrate the different healthcare data sources (*'D5: Consider Data at Multiple Abstraction Levels'*), and handle unexpected behaviour when an unpredictable event occurs (*'D4: Break the Glass'*) or when a deadline will be violated [158] (*'D2: Value the Infrequent Behaviour'*), even being able to perform immediate decision making while analyzing the data in real time (e.g., streaming process mining) [159].

### 4.10. C10: Evolve in symbiosis with the developments in the healthcare domain (D1, D3, D7, D10)

Healthcare is a domain that continuously develops due to

innovations in, amongst others, medicines, medical procedures and technologies. In order to support healthcare professionals on an ongoing basis, PM4H methods should evolve in symbiosis with the developments in healthcare. For instance: treatment processes are rapidly being adapted based on emerging scientific evidence [2]. A clear illustration is the COVID-19 treatment process, which rapidly changed due to the accumulation of new insights in the virus and its treatment [133,160]. PM4H should be able to cope with these circumstances. Moreover, PM4H can support the adaptation of treatment processes by, e.g., enabling the efficient comparison of processes with respect to the clinical outcomes they generate.

Another important evolution in healthcare is the prominent presence of personalised medicine, which implies that medical treatments are increasingly tailored towards the needs of each individual patient [50,161]. In order to support personalised medicine, PM4H can develop techniques to efficiently assess the suitability of a particular treatment process for a patient with a specific profile. For instance, Valero-Ramon et al. [89] recently proposed an approach to discover dynamic risk models for patients suffering from chronic diseases based on sensor data. These models can be leveraged to customise treatments based on a patient's unique behaviour [89]. When process data can be enriched with outcome and cost data, PM4H also has the potential to study the effect of personalised treatment processes compared to standard practice in detail.

## 5. Concluding remarks

The goal of this paper is to support the PM4H domain by conveying a shared perspective regarding the distinguishing characteristics of healthcare processes and the associated key challenges of PM4H. In particular, ten distinguishing characteristics and ten challenges are outlined, the latter requiring specific attention of the PM4H community. In order to tackle the key challenges, close collaboration is required between experts from various domains. In that respect, the symbiosis between data- and process-related expertise on the one hand and clinical expertise on the other is essential. Within a research context, close interaction is required between research fields including medicine, medical informatics, computer science, and business process management. To make PM4H flourish, researchers from these various fields need to join forces in a spirit of co-creation. In this way, they may be able to develop innovative process mining methods that maximally support clinicians and decision-makers to manage and improve real-world healthcare processes in an evidence-based way.

While this paper shapes the identity of PM4H, there is no intention to position PM4H as an isolated domain. On the contrary: as a relatively young research field, the PM4H community should actively connect to and learn from more mature fields such as artificial intelligence, data mining and machine learning [162,163]. The strengths of PM4H will be based on intelligently combining insights from various research areas. This intersection of disciplines will increasingly involve research on big data, the Internet of Things and deep learning [164–166]. Translating concepts and techniques from various other research fields to the PM4H setting will present continuous research challenges. Hence, there is much to be done and much to be achieved.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work is partially supported by ANID FONDECYT 1220202, Dirección de Investigación de la Vicerrectoría de Investigación de la Pontificia Universidad Católica de Chile - PUENTE [Grant No. 026/2021]; and Agencia Nacional de Investigación y Desarrollo [Grant Nos. ANID-PFCHA/Doctorado Nacional/2019–21190116, ANID-PFCHA/Doctorado Nacional/2020–21201411]. With regard to the co-author Hilda Klasky, this manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

## References

[1] R.E. Herzlinger, Why innovation in health care is so hard, Harvard Bus. Rev. 84 (5) (2006) 58–66.
[2] N. Peiffer-Smadja, J.-C. Lucet, G. Bendjelloul, L. Bouadma, S. Gerard, C. Choquet, S. Jacques, A. Khalil, P. Maisani, E. Casalino, et al., Challenges and issues about organizing a hospital to respond to the COVID-19 outbreak: experience from a french reference centre, Clin. Microbiol. Infect. 26 (6) (2020) 669–672.
[3] E. Coiera, E.S. Hovenga, Building a sustainable health system, Yearbook Med. Inform. 16 (1) (2007) 11–18.
[4] P.R. Harper, A framework for operational modelling of hospital resources, Health Care Manage. Sci. 5 (3) (2002) 165–173.
[5] R.S. Mans, W.M.P. van der Aalst, R.J.B. Vanwersch, Process mining in healthcare: evaluating and exploiting operational healthcare processes, Springer International Publishing, Heidelberg, 2015.
[6] R. Lenz, M. Peleg, and M. Reichert, "Healthcare process support: achievements, challenges, current research," International Journal of Knowledge-Based Organizations, vol. 2, no. 4, 2012.
[7] N. Martin, J. De Weerdt, C. Fernández-Llatas, A. Gal, R. Gatta, G. Ibá nez, O. Johnson, F. Mannhardt, L. Marco-Ruiz, S. Mertens, et al., "Recommendations for enhancing the usability and understandability of process mining in healthcare," Artificial Intelligence in Medicine, vol. 109, 2020.
[8] W.M.P. van der Aalst, Process mining: data science in action, Springer-Verlag, Berlin Heidelberg, Heidelberg, 2016.
[9] Z. Huang, W. Dong, L. Ji, L. Yin, H. Duan, On local anomaly detection and analysis for clinical pathways, Artif. Intell. Med. 65 (3) (2015) 167–177.
[10] E. Rojas, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Process mining in healthcare: a literature review, J. Biomed. Inform. 61 (2016) 224–236.
[11] 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 Trans. Knowl. Data Eng. 31 (4) (2018) 686–705.
[12] J. Carmona, B. van Dongen, A. Solti, M. Weidlich, Conformance checking, Springer International Publishing, Heidelberg, 2018.
[13] C. Cabanillas, L. Ackermann, S. Schönig, C. Sturm, J. Mendling, The RALph miner for automated discovery and verification of resource-aware process models, Software and Systems Modeling 19 (6) (2020) 1415–1441.
[14] P. Barach, J. Johnson, Understanding the complexity of redesigning care around the clinical microsystem, BMJ Quality & Safety 15 (suppl 1) (2006) i10–i16.
[15] T. Bodenheimer, C. Sinsky, From triple to quadruple aim: care of the patient requires care of the provider, The Annals of Family Medicine 12 (6) (2014) 573–576.
[16] D.M. Berwick, T.W. Nolan, J. Whittington, The triple aim: care, health, and cost, Health Aff. 27 (3) (2008) 759–769.
[17] M. Kerremans, S. Searle, T. Srivastava, and K. Iijima, "Market Guide for Process Mining," tech. rep., Gartner, 2020.
[18] Á. Rebuge, D.R. Ferreira, Business process analysis in healthcare environments: a methodology based on process mining, Information Systems 37 (2) (2012) 99–116.
[19] C. Di Ciccio, A. Marrella, A. Russo, Knowledge-intensive processes: characteristics, requirements and analysis of contemporary approaches, Journal on Data Semantics 4 (1) (2015) 29–57.
[20] S. Mertens, F. Gailly, G. Poels, Towards a decision-aware declarative process modeling language for knowledge-intensive processes, Expert Syst. Appl. 87 (2017) 316–334.
[21] T.G. Erdogan, A. Tarhan, Systematic mapping of process mining studies in healthcare, IEEE Access 6 (2018) 24543–24567.
[22] M. Dumas, M. Rosa, J. Mendling, H.A. Reijers, Fundamentals of Business Process Management, Springer, Berlin Heidelberg, 2018.
[23] G.T. Jun, J. Ward, Z. Morris, J. Clarkson, Health care process modelling: which method when? Int. J. Qual. Health Care 21 (3) (2009) 214–224.
[24] H. Scheuerlein, F. Rauchfuss, Y. Dittmar, R. Molle, T. Lehmann, N. Pienkos, U. Settmacher, New methods for clinical pathways—business process modeling notation (BPMN) and tangible business process modeling (t.BPM), Langenbeck's Archives of Surgery 397 (5) (2012) 755–761.
[25] F. Mannhardt, Multi-perspective process mining, CEUR Workshop Proceedings 2196 (2018) 41–45.

[26] F. Mannhardt, D. Blind, Analyzing the trajectories of patients with sepsis using process mining, CEUR Workshop Proceedings 1859 (2017) 72–80.
[27] S. Remy, L. Pufahl, J.P. Sachs, E. Böttinger, M. Weske, Event log generation in a health system: a case study, Lect. Notes Comput. Sci. 12168 (2020) 505–522.
[28] L. Leiva, J. Munoz-Gama, J. Salas-Morales, V. Gálvez, W.L.J. Lee, R. de la Fuente, R. Fuentes, M. Sepúlveda, POMElog: generating event logs from unplugged processes, CEUR Workshop Proceedings 2420 (2019) 189–193.
[29] J. Munoz-Gama, V. Galvez, R. de la Fuente, M. Sepúlveda, and R. Fuentes, "Interactive process mining for medical training," in Interactive Process Mining in Healthcare (C. Fernandez-Llatas, ed.), pp. 233–242, Springer International Publishing, 2021.
[30] C. Alvarez, E. Rojas, M. Arias, J. Munoz-Gama, M. Sepúlveda, V. Herskovic, D. Capurro, Discovering role interaction models in the emergency room using process mining, J. Biomed. Inform. 78 (2018) 60–77.
[31] M. Song, W.M.P. van der Aalst, Towards comprehensive support for organizational mining, Decis. Support Syst. 46 (1) (2008) 300–317.
[32] W.M.P. van der Aalst, H.A. Reijers, M. Song, Discovering social networks from event logs, Computer Supported Cooperative Work (CSCW) 14 (6) (2005) 549–593.
[33] 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.
[34] S. Suriadi, M.T. Wynn, J. Xu, W.M.P. van der Aalst, A.H. ter Hofstede, Discovering work prioritisation patterns from event logs, Decis. Support Syst. 100 (2017) 77–92.
[35] J. De Weerdt, F. Caron, J. Vanthienen, B. Baesens, Getting a grasp on clinical pathway data: an approach based on process mining, Lect. Notes Comput. Sci. 7769 (2013) 22–35.
[36] J. Munoz-Gama, "Conformance checking," in Encyclopedia of Big Data Technologies (S. Sakr and A.Y. Zomaya, eds.), Springer, 2019.
[37] T. Conca, C. Saint-Pierre, V. Herskovic, M. Sepúlveda, D. Capurro, F. Prieto, C. Fernandez-Llatas, Multidisciplinary collaboration in the treatment of patients with type 2 diabetes in primary care: analysis using process mining, Journal of Medical Internet Research 20 (4) (2018) e127.
[38] D. Fahland, W.M.P. van der Aalst, Model repair - aligning process models to reality, Information Systems 47 (2015) 220–243.
[39] M. Rovani, F.M. Maggi, M. de Leoni, W.M.P. van der Aalst, Declarative process mining in healthcare, Expert Syst. Appl. 42 (23) (2015) 9236–9251.
[40] C. Combi, F. Galetto, H.C. Nakawala, G. Pozzi, F. Zerbato, Enriching surgical process models by BPMN extensions for temporal durations, in: Proceedings of the 36th Annual ACM Symposium on Applied Computing, 2021, pp. 586–593.
[41] E. Batista and A. Solanas, "Process mining in healthcare: a systematic review," Proceedings of the 9th International Conference on Information, Intelligence, Systems and Applications (IISA), pp. 1–6, 2018.
[42] F. Taymouri, M.L. Rosa, M. Dumas, F.M. Maggi, Business process variant analysis: Survey and classification, Knowl.-Based Syst. 211 (2021) 106557.
[43] F. Caron, J. Vanthienen, K. Vanhaecht, E.V. Limbergen, J. Deweerdt, B. Baesens, A process mining-based investigation of adverse events in care processes, Health Information Management Journal 43 (1) (2014) 16–25.
[44] W. Yang and Q. Su, "Process mining for clinical pathway: Literature review and future directions," 2014 11th International Conference on Service Systems and Service Management (ICSSSM), pp. 1–5, 2014.
[45] F. Caron, J. Vanthienen, K. Vanhaecht, E.V. Limbergen, J. De Weerdt, B. Baesens, Monitoring care processes in the gynecologic oncology department, Comput. Biol. Med. 44 (2014) 88–96.
[46] R. Andrews, M.T. Wynn, K. Vallmuur, A.H.M. ter Hofstede, and E. Bosley, "A comparative process mining analysis of road trauma patient pathways," International Journal of Environmental Research and Public Health, vol. 17, no. 10, 2020.
[47] A.B. Jensen, P.L. Moseley, T.I. Oprea, S.G. Ellesøe, R. Eriksson, H. Schmock, P. B. Jensen, L.J. Jensen, S. Brunak, "Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients,", Nature Communications 5 (1) (2014) 4022.
[48] G. Kusuma, A.P. Kurniati, E. Rojas, C.D. McInerney, C.P. Gale, O.A. Johnson, Process mining of disease trajectories: A literature review, Studies in Health Technology and Informatics 281 (2021) 457–461.
[49] H. De Oliveira, M. Prodel, L. Lamarsalle, M. Inada-Kim, K. Ajayi, J. Wilkins, S. Sekelj, S. Beecroft, S. Snow, R. Slater, et al., bow-tie optimal pathway discovery analysis of sepsis hospital admissions using the hospital episode statistics database in england, JAMIA open 3 (3) (2020) 439–448.
[50] I.R. König, O. Fuchs, G. Hansen, E. von Mutius, and M.V. Kopp, "What is precision medicine?," European Respiratory Journal, vol. 50, no. 4, 2017.
[51] D.L. Sackett, W.M.C. Rosenberg, M.J.A. Gray, B.R. Haynes, S.W. Richardson, Evidence based medicine: what it is and what it isn't, BMJ 312 (7023) (1996) 71–72.
[52] E. Rojas, A. Cifuentes, A. Burattin, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Analysis of emergency room episodes duration through process mining, Lecture Notes in Business Information Processing 342 (2018) 251–263.
[53] J.F.J. Vos, A. Boonstra, A. Kooistra, M. Seelen, M. van Offenbeek, The influence of electronic health record use on collaboration among medical specialties, BMC Health Services Research 20 (1) (2020), 676-676.
[54] R.S. Mans, M. Schonenberg, M. Song, W. Van der Aalst, P. Bakker, Process mining in healthcare, in: in Proceedings of the International Conference on Health Informatics, 2008, pp. 118–125.
[55] S. Pereira Detro, E.A.P. Santos, H. Panetto, E.D. Loures, M. Lezoche, C. Cabral Moro Barra, Applying process mining and semantic reasoning for process model customisation in healthcare, Enterprise Information Systems 14 (7) (2020) 983–1009.
[56] C.P. McLaughlin, A.D. Kaluzny, Continuous quality improvement in health care, Jones and Bartlett Publishers, Sudbury, 2004.
[57] J. De Weerdt, "Trace clustering," Encyclopedia of Big Data Technologies, pp. 1706–1711, 2019.
[58] S. Smirnov, H.A. Reijers, M. Weske, From fine-grained to abstract process models: A semantic approach, Information Systems 37 (8) (2012) 784–797.
[59] G. Leonardi, M. Striani, S. Quaglini, A. Cavallini, S. Montani, Leveraging semantic labels for multi-level abstraction in medical process mining and trace comparison, J. Biomed. Inform. 83 (2018) 10–24.
[60] S. Alter, Theory of workarounds, Communications of the Association for Information Systems 34 (2014) 55.
[61] I. Beerepoot, X. Lu, I. van de Weerd, and H.A. Reijers, "Seeing the signs of workarounds: a mixed-methods approach to the detection of nurses' process deviations," Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS), pp. 1–10, 2021.
[62] L. Vanbrabant, N. Martin, K. Ramaekers, K. Braekers, Quality of input data in emergency department simulations: Framework and assessment techniques, Simul. Model. Pract. Theory 91 (2019) 83–101.
[63] I. Beerepoot, I. van de Weerd, H.A. Reijers, The potential of workarounds for improving processes, Lecture Notes in Business Information Processing 362 (2019) 338–350.
[64] M. Robin DiMatteo, P.J. Giordani, H.S. Lepper, and T.W. Croghan, "Patient adherence and medical treatment outcomes: a meta-analysis," Medical Care, vol. 40, no. 9, 2002.
[65] A. Ten Teije, M. Marcos, M. Balser, J. van Croonenborg, C. Duelli, F. van Harmelen, P. Lucas, S. Miksch, W. Reif, K. Rosenbrand, et al., Improving medical protocols by formal methods, Artif. Intell. Med. 36 (3) (2006) 193–209.
[66] D. Wang, M. Peleg, S.W. Tu, A.A. Boxwala, R.A. Greenes, V.L. Patel, E. H. Shortliffe, Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines: a literature review of guideline representation models, Int. J. Med. Informatics 68 (1–3) (2002) 59–70.
[67] K.N. Lohr, M.J. Field, et al., Clinical practice guidelines: directions for a new program, National Academies Press, 1990.
[68] R. Gatta, M. Vallati, C. Fernandez-Llatas, A. Martinez-Millana, S. Orini, L. Sacchi, J. Lenkowicz, M. Marcos, J. Munoz-Gama, M.A. Cuendet, et al., "What role can process mining play in recurrent clinical guidelines issues? a position paper," International Journal of Environmental Research and Public Health, vol. 17, no. 18, 2020.
[69] R. Gatta, M. Vallati, C. Fernández-Llatas, A. Martinez-Millana, S. Orini, L. Sacchi, J. Lenkowicz, M. Marcos, J. Munoz-Gama, M.A. Cuendet, B.D. Bari, L. Marco-Ruiz, A. Stefanini, M. Castellano, Clinical guidelines: a crossroad of many research areas. Challenges and opportunities in process mining for healthcare, Lecture Notes in Business Information Processing 362 (2019) 545–556.
[70] S. Montani, G. Leonardi, S. Quaglini, A. Cavallini, G. Micieli, A knowledge-intensive approach to process similarity calculation, Expert Syst. Appl. 42 (9) (2015) 4207–4215.
[71] R. Bergmann, Y. Gil, Similarity assessment and efficient retrieval of semantic workflows, Information Systems 40 (2014) 115–127.
[72] M. La Rosa, M. Dumas, R. Uba, R. Dijkman, Business process model merging: An approach to business process consolidation, ACM Transactions on Software Engineering and Methodology 22 (2) (2013) 1–42.
[73] A. Goderis, P. Li, C. Goble, Workflow discovery: the problem, a case study from e-science and a graph-based solution, IEEE International Conference on Web Services ICWS'06 (2006) 312–319.
[74] R. de la Fuente, R. Fuentes, J. Munoz-Gama, J. Dagnino, M. Sepúlveda, Delphi method to achieve clinical consensus for a BPMN representation of the central venous access placement for training purposes, International Journal of Environmental Research and Public Health 17 (11) (2020).
[75] R. de la Fuente, E. Kattan, J. Munoz-Gama, I. Puente, M. Navarrete, C. Kychenthal, R. Fuentes, S. Bravo, V. Galvez, M. Sepúlveda, Development of a comprehensive percutaneous dilatational tracheostomy process model for procedural training: a delphi-based experts consensus, Acta Anaesthesiol. Scand. 65 (2) (2021) 244–256.
[76] S. Timmermans, M. Berg, Standardization in action: achieving local universality through medical protocols, Soc. Stud. Sci. 27 (2) (1997) 273–305.
[77] L.O. Karlsson, S. Nilsson, M. Bång, L. Nilsson, E. Charitakis, M. Janzon, A clinical decision support tool for improving adherence to guidelines on anticoagulant therapy in patients with atrial fibrillation at risk of stroke: a cluster-randomized trial in a swedish primary care setting (the CDS-AF study), PLoS Medicine 15 (3) (2018) 1–17.
[78] M. Peleg, Computer-interpretable clinical guidelines: a methodological review, J. Biomed. Inform. 46 (4) (2013) 744–763.
[79] D. Roulin, M. Muradbegovic, V. Addor, C. Blanc, N. Demartines, M. Hübner, Enhanced recovery after elective colorectal surgery-reasons for non-compliance with the protocol, Digestive Surgery 34 (3) (2017) 220–226.
[80] V. van Andel, I. Beerepoot, X. Lu, I. van de Weerd, H.A. Reijers, DEUCE: a methodology for detecting unauthorized access of electronic health records using process mining, European Conference on Information Systems (ECIS) Research Papers (2021).
[81] 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), 2016, pp. 291–297.

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

[83] R. Williams, E. Rojas, N. Peek, O.A. Johnson, Process mining in primary care: a literature review, Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth 247 (2018) 376–380.

[84] N.F. Farid, M. de Kamps, and O.A. Johnson, "Process mining in frail elderly care: a literature review," Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies, vol. 5, pp. 332–339, 2019.

[85] H.B. Klasky and O. Ozmen, "Process mining in healthcare - a case study for the corporate data warehouse of the veterans affairs office," tech. rep., Office of Scientific and Technical Information (OSTI), 2019.

[86] S.J. van Zelst, F. Mannhardt, M. de Leoni, A. Koschmider, Event abstraction in process mining: literature review and taxonomy, Granular Computing 6 (2020) 719–736.

[87] C. Günther, A. Rozinat, W.M.P. van der Aalst, and K. van Uden, "Monitoring deployed application usage with process mining," BPM reports, 2008.

[88] S.N. Araghi, F. Fontanili, E. Lamine, N. Salatgé, J. Lesbegueries, S.R. Pouyade, and F. Bénaben, "Evaluating the process capability ratio of patients' pathways by the application of process mining, SPC and RTLS", Proceedings of the 12th International Conference on Health Informatics (HEALTHINF), vol. 5, pp. 302–309, 2019.

[89] Z. Valero-Ramon, C. Fernández-Llatas, B. Valdivieso, V. Traver, Dynamic models supporting personalised chronic disease management through healthcare sensors with interactive process mining, Sensors 20 (18) (2020).

[90] S. Suriadi, R. Mans, M.T. Wynn, A. Partington, J. Karnon, Measuring patient flow variations: a cross-organisational process mining approach, Lecture Notes in Business Information Processing 181 (2014) 43–58.

[91] J.T. Marquinez, A. Sauré, A. Cataldo, J.-C. Ferrer, Identifying proactive ICU patient admission, transfer and diversion policies in a public-private hospital network, Eur. J. Oper. Res. 295 (2021) 306–320.

[92] L. Cao, Data science: a comprehensive overview, ACM Computing Surveys (CSUR) 50 (3) (2017) 1–42.

[93] F. Veit, J. Geyer-Klingeberg, J. Madrzak, M. Haug, J. Thomson, The proactive insights engine: Process mining meets machine learning and artificial intelligence, CEUR Workshop Proceedings 1920 (2017) 1–5.

[94] E. Helm, A.M. Lin, D. Baumgartner, A.C. Lin, J. Küng, Towards the use of standardized terms in clinical case studies for process mining in healthcare, International Journal of Environmental Research and Public Health 17 (4) (2020) 1348.

[95] M. Arias, E. Rojas, S. Aguirre, F. Cornejo, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Mapping the patient's journey in healthcare through process mining, International Journal of Environmental Research and Public Health 17 (18) (2020).

[96] M. Berghout, J. Van Exel, L. Leensvaart, J.M. Cramm, Healthcare professionals' views on patient-centered care in hospitals, BMC Health Services Research 15 (1) (2015) 1–13.

[97] A.I. Poon, J.J. Sung, Opening the black box of AI-Medicine, J. Gastroenterol. Hepatol. 36 (3) (2021) 581–584.

[98] M. Jans, P. Soffer, T. Jouck, Building a valuable event log for process mining: an experimental exploration of a guided process, Enterprise Information Systems 13 (5) (2019) 601–630.

[99] C. Fernandez-Llatas, ed., Interactive Process Mining in Healthcare. Springer International Publishing, 2020.

[100] A. Latoszek-Berendsen, H. Tange, H. Van Den Herik, A. Hasman, From clinical practice guidelines to computer-interpretable guidelines, Methods of information in medicine 49 (06) (2010) 550–570.

[101] M. Arias, E. Rojas, J. Munoz-Gama, M. Sepúlveda, A framework for recommending resource allocation based on process mining, Lecture Notes in Business Information Processing 256 (2016) 458–470.

[102] A. Pika, M.T. Wynn, S. Budiono, A.H. ter Hofstede, W.M.P. van der Aalst, and H.A. Reijers, "Privacy-preserving process mining in healthcare," International Journal of Environmental Research and Public Health, vol. 17, no. 5, 2020.

[103] C. Safran, M. Bloomrosen, W.E. Hammond, S. Labkoff, S. Markel-Fox, P.C. Tang, D.E. Detmer, Toward a national framework for the secondary use of health data: an american medical informatics association white paper, J. Am. Med. Inform. Assoc. 14 (1) (2007) 1–9.

[104] N. Martin, J. Bergs, Patient flow data registration: a key barrier to the data-driven and proactive management of an emergency department, International Emergency Nursing 53 (2020).

[105] R.P.J.C. Bose, R.S. Mans, and W.M.P. van der Aalst, "Wanna improve process mining results?," Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining (CIDM), pp. 127–134, 2013.

[106] N. Martin, "Data quality in process mining," in Interactive Process Mining in Healthcare (C. Fernandez-Llatas, ed.), pp. 53–79, Springer International Publishing, 2021.

[107] B. Djulbegovic, G.H. Guyatt, Progress in evidence-based medicine: a quarter century on, The Lancet 390 (10092) (2017) 415–423.

[108] P. Dimitri, "The evolution of evidence based clinical medicine," in Practical pediatric urology (P. Godbole and M. Wilcox, D.T. adn Koyle, eds.), pp. 1–15, Cham: Springer, 2021.

[109] B.M. Silva, J.J. Rodrigues, I. de la Torre Díez, M. López-Coronado, K. Saleem, Mobile-health: a review of current state in 2015, J. Biomed. Inform. 56 (2015) 265–272.

[110] C. Rathert, M.D. Wyrwich, S.A. Boren, Patient-centered care and outcomes: a systematic review of the literature, Medical Care Research and Review 70 (4) (2013) 351–379.

[111] D. Miller, C.S. Gray, K. Kuluski, C. Cott, Patient-centered care and patient-reported measures: let's look before we leap, The Patient-Patient-Centered Outcomes Research 8 (4) (2015) 293–299.

[112] M.L. van Eck, X. Lu, S.J.J. Leemans, W.M.P. van der Aalst, PM$^2$: a process mining project methodology, Lect. Notes Comput. Sci. 9097 (2015) 297–313.

[113] S. Aguirre, C. Parra, M. Sepúlveda, Methodological proposal for process mining projects, International Journal of Business Process Integration and Management 8 (2) (2017) 102–113.

[114] J.J. Maldonado, R. Palta, J. Vazquez, J.L. Bermeo, M. Pérez-Sanagustín, J. Munoz-Gama, Exploring differences in how learners navigate in MOOCs based on self-regulated learning and learning styles: a process mining approach, in: Proceedings of the Latin American Computing Conference CLEI, 2016, pp. 1–12.

[115] O.A. Johnson, T.B. Dhafari, A. Kurniati, F. Fox, E. Rojas, The ClearPath method for care pathway process mining and simulation, Lecture Notes in Business Information Processing 342 (2018) 239–250.

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

[117] 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 (3) (2017) 302.

[118] A. Rozinat, A.K.A. de Medeiros, C.W. Günther, A.J.M.M. Weijters, W.M.P. van der Aalst, The need for a process mining evaluation framework in research and practice, Lect. Notes Comput. Sci. 4928 (2007) 84–89.

[119] M.J. Page, J.E. McKenzie, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M.M. Lalu, T. Li, E.W. Loder, E. Mayo-Wilson, S. McDonald, L.A. McGuinness, L.A. Stewart, J. Thomas, A.C. Tricco, V.A. Welch, P. Whiting, D. Moher, The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, BMJ 372 (2020) 2021.

[120] A.R.C. Maita, L.C. Martins, C.R. Lopez Paz, L. Rafferty, P.C. Hung, S.M. Peres, M. Fantinato, A systematic mapping study of process mining, Enterprise Information Systems 12 (5) (2018) 505–549.

[121] N. Martin, B. Depaire, A. Caris, The use of process mining in business process simulation model construction, Business & Information Systems Engineering 58 (1) (2016) 73–87.

[122] A.E. Márquez-Chamorro, M. Resinas, A. Ruiz-Cortés, Predictive monitoring of business processes: a survey, IEEE Trans. Serv. Comput. 11 (6) (2018) 962–977.

[123] W. Chomyat and W. Premchaiswadi, "Process mining on medical treatment history using conformance checking," Proceedings of the 14th International Conference on ICT and Knowledge Engineering (ICT KE), pp. 77–83, 2016.

[124] 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, Postgrad. Med. J. 96 (1135) (2020) 250–256.

[125] C. Rinner, E. Helm, R. Dunkl, H. Kittler, S. Rinderle-Ma, An application of process mining in the context of melanoma surveillance using time boxing, Lecture Notes in Business Information Processing 342 (2018) 175–186.

[126] M.R. Dallagassa, C. dos Santos Garcia, E.E. Scalabrin, S.O. Ioshii, D.R. Carvalho, Opportunities and challenges for applying process mining in healthcare: a systematic mapping study, Journal of Ambient Intelligence and Humanized Computing (2021) 1–18.

[127] A. Stefanini, D. Aloini, E. Benevento, R. Dulmin, V. Mininno, Performance analysis in emergency departments: a data-driven approach, Measuring Business Excellence 22 (2) (2018) 130–145.

[128] A. Stefanini, D. Aloini, E. Benevento, R. Dulmin, V. Mininno, A data-driven methodology for supporting resource planning of health services, Socio-Economic Planning Sciences 70 (2020) 100744.

[129] E. Benevento, D. Aloini, N. Squicciarini, R. Dulmin, V. Mininno, Queue-based features for dynamic waiting time prediction in emergency department, Measuring Business Excellence 23 (4) (2019) 458–471.

[130] R.P.J.C. Bose, W.M.P. van der Aalst, I. Žliobaitė, M. Pechenizkiy, Handling concept drift in process mining, Lect. Notes Comput. Sci. 6741 (2011) 391–405.

[131] A. Maaradji, M. Dumas, M. La Rosa, A. Ostovar, Detecting sudden and gradual drifts in business processes from execution traces, IEEE Trans. Knowl. Data Eng. 29 (10) (2017) 2140–2154.

[132] A. Yeshchenko, C. Di Ciccio, J. Mendling, A. Polyvyanyy, Comprehensive process drift detection with visual analytics, Lect. Notes Comput. Sci. 11788 (2019) 119–135.

[133] P. Chibber, S.A. Haq, I. Ahmed, N.I. Andrabi, G. Singh, Advances in the possible treatment of COVID-19: a review, Eur. J. Pharmacol. 883 (2020).

[134] C.J. Turner, A. Tiwari, R. Olaiya, Y. Xu, Process mining: from theory to practice, Business Process Management Journal 18 (3) (2012) 493–512.

[135] S.J. Leemans, E. Poppe, and M.T. Wynn, "Directly follows-based process mining: Exploration & a case study," Proceedings of the International Conference on Process Mining (ICPM), vol. 25–32, 2019.

[136] M. Weske, Business Process Management - Concepts, Languages, Architectures, Third Edition, Springer-Verlag, Berlin Heidelberg, 2019.

[137] F. Gonzalez-Lopez, L. Pufahl, J. Munoz-Gama, V. Herskovic, M. Sepúlveda, Case model landscapes: toward an improved representation of knowledge-intensive processes using the fCM-language, Software and Systems Modeling (2021).

[138] W.M.P. van der Aalst, A practitioner's guide to process mining: Limitations of the directly-follows graph, Procedia Computer Science 164 (2019) 321–328.

[139] R. Andrews, C. van Dun, M. Wynn, W. Kratsch, M. Röglinger, A. ter Hofstede, Quality-informed semi-automated event log generation for process mining, Decis. Support Syst. 132 (2020) 113265.
[140] D.A. Fischer, K. Goel, R. Andrews, C.G.J. van Dun, M.T. Wynn, M. Röglinger, Enhancing event log quality: Detecting and quantifying timestamp imperfections, Lect. Notes Comput. Sci. 12168 (2020) 309–326.
[141] 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 (2019) 226–238.
[142] 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.
[143] K. Diba, K. Batoulis, M. Weidlich, M. Weske, Extraction, correlation, and abstraction of event data for process mining, WIREs Data Mining and Knowledge Discovery 10 (3) (2020) e1346.
[144] 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.
[145] W.M.P. van der Aalst, M. Bichler, A. Heinzl, Responsible data science, Business & Information Systems Engineering 59 (5) (2017) 311–313.
[146] W.M.P. van der Aalst, A. Adriansyah, A.K.A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, R.P.J.C. Bose, P. van den Brand, R. Brandtjen, J. Buijs, A. Burattin, J. Carmona, et al., Process mining manifesto, Lecture Notes in Business Information Processing 99 (2012) 169–194.
[147] M. Rafiei, W.M.P. van der Aalst, Privacy-preserving data publishing in process mining, Lecture Notes in Business Information Processing 392 (2020) 122–138.
[148] S. Nu nez von Voigt, S.A. Fahrenkrog-Petersen, D. Janssen, A. Koschmider, F. Tschorsch, F. Mannhardt, O. Landsiedel, and M. Weidlich, "Quantifying the re-identification risk of event logs for process mining," Lecture Notes in Computer Science, vol. 12127, pp. 252–267, 2020.
[149] F. Rapport, P. Hibbert, M. Baysari, J. Long, R. Seah, W. Zheng, C. Jones, K. Preece, J. Braithwaite, What do patients really want? An in-depth examination of patient experience in four australian hospitals, BMC Health Services Research 19 (1) (2019) 1–9.
[150] J.D. Sonis, E.L. Aaronson, R.Y. Lee, L.L. Philpotts, B.A. White, Emergency department patient experience: a systematic review of the literature, Journal of Patient Experience 5 (2) (2018) 101–106.
[151] J.S. McCullough, The adoption of hospital information systems, Health Econ. 17 (5) (2008) 649–664.
[152] P.L. Reichertz, Hospital information systems–past, present, future, Int. J. Med. Informatics 75 (3–4) (2006) 282–299.
[153] K.A. Kuhn, D.A. Giuse, From hospital information systems to health information systems-problems, challenges, perspectives, Yearbook of Medical Informatics 10 (1) (2001) 63–76.
[154] M. Dumas, J. Recker, M. Weske, Management and engineering of process-aware information systems: introduction to the special issue, Information Systems 37 (2) (2012) 77–79.
[155] M. Dumas, W.M.P. van der Aalst, and A.H.M. ter Hofstede, eds., Process-Aware Information Systems: Bridging People and Software Through Process Technology. Wiley, 2005.
[156] M.J. Ball, Hospital information systems: perspectives on problems and prospects, 1979 and 2002, Int. J. Med. Informatics 69 (2–3) (2003) 83–89.
[157] M. Reichert, B. Weber, Enabling flexibility in process-aware information systems: challenges, methods, technologies, Springer-Verlag, Berlin Heidelberg, 2012.
[158] W.M.P. van der Aalst, M. Rosemann, M. Dumas, Deadline-based escalation in process-aware information systems, Decis. Support Syst. 43 (2) (2007) 492–511.
[159] A. Burattin, Streaming process discovery and conformance checking, Encyclopedia of Big Data Technologies (2018) 1–8.
[160] P. Tarighi, S. Eftekhari, M. Chizari, M. Sabernavaei, D. Jafari, P. Mirzabeigi, A review of potential suggested drugs for coronavirus disease (covid-19) treatment, Eur. J. Pharmacol. (2021) 173890.
[161] H.K. Brittain, R. Scott, E. Thomas, The rise of the genome and personalised medicine, Clinical Medicine 17 (6) (2017) 545–551.
[162] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, Y. Wang, Artificial intelligence in healthcare: past, present and future, Stroke and Vascular Neurology 2 (4) (2017) 230–243.
[163] N. Jothi, W. Husain, et al., Data mining in healthcare–a review, Procedia Computer Science 72 (2015) 306–313.
[164] F. Chen, P. Deng, J. Wan, D. Zhang, A.V. Vasilakos, X. Rong, Data mining for the internet of things: literature review and challenges, Int. J. Distrib. Sens. Netw. 11 (8) (2015).
[165] G. Manogaran, D. Lopez, A survey of big data architectures and machine learning algorithms in healthcare, Int. J. Biomed. Eng. Technol. 25 (2–4) (2017) 182–211.
[166] R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, Deep learning for healthcare: review, opportunities and challenges, Briefings in Bioinformatics 19 (6) (2017) 1236–1246.