

Using Process Mining in Healthcare

Niels Martin^{1,2}(⊠), Nils Wittig³, and Jorge Munoz-Gama⁴

¹ Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium niels.martin@uhasselt.be

² Research Foundation Flanders (FWO), Egmontstraat 5, 1000 Brussels, Belgium ³ KMS Vertrieb und Services AG, Inselkammerstraße 1,

82008 Unterhaching, Germany

nils.wittig@kms.ag

⁴ Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860,

7820436 Macul, Chile

jmun@uc.cl

Abstract. This chapter introduces a specific application domain of process mining: healthcare. Healthcare is a very promising domain for process mining given the significant societal value that can be generated by supporting process improvement in a data-driven way. Within a healthcare organisation, a wide variety of processes is being executed, many of them being highly complex due to their loosely-structured and knowledge-intensive nature. Consequently, performing process mining in healthcare is challenging, but can generate significant societal impact. To provide more insights in process mining in healthcare, this chapter first provides an overview of healthcare processes and healthcare process data, as well as their particularities compared to other domains. Afterwards, an overview of common use cases in process mining in healthcare research is presented, as well as insights from a real-life case study. Subsequently, an overview of open challenges to ensure a widespread adoption of process mining in healthcare is provided. By tackling these challenges, process mining will become able to fully play its role to support evidence-based process improvement in healthcare and, hence, contribute to shaping the best possible care for patients in a way that is sustainable in the long run.

Keywords: Process mining \cdot Healthcare \cdot Evidence-based process improvement

1 Introduction

The prior chapters of this book introduced various process mining topics. In contrast to these preceding chapters, this chapter focuses on introducing a specific application domain of process mining. In particular, this chapter focuses on *healthcare*. In process mining research, healthcare illustrations are often used to demonstrate new techniques, or a healthcare problem is the starting point

of the research project altogether [55]. This can be, at least partly, explained by the great societal value related to efforts to improve the healthcare system. In many countries, the long-term sustainability of the healthcare system is an important societal issue due to trends such as the increasing life expectancy, and the raising prevalence of chronic diseases [29]. Improvements in terms of healthcare processes is an indispensable piece of the puzzle to sustain the healthcare system, while continuously improving the quality of care delivered to the patient.

Within the healthcare domain, many different processes are being performed in a wide variety of healthcare organisations. Many processes in healthcare are complex as they are loosely-framed and knowledge-intensive [20,55,58]. While the former indicates that healthcare processes can typically be executed in a large number of distinct ways [58], the latter indicates that the trajectory that is followed strongly depends upon complex decisions made by knowledge workers such as physicians and nurses [20]. These healthcare processes are increasingly being supported by health information systems [53], which capture data about the real-life execution of a process in their databases. This data can be leveraged to compose an event log, the key input for process mining [55].

There has been a steady growth in research interest on process mining in healthcare in recent years [17]. Despite the great potential of process mining to support process improvement in healthcare and the increasing number of methods specifically designed for the healthcare context, the systematic uptake of process mining in healthcare organisations outside the research context is still fairly limited [55]. Hence, there are still challenges ahead that need to be overcome, which is consistent with the fact that process mining in healthcare is a rather young research area. Moreover, healthcare is a highly dynamic field as processes change due to advances in, for instance, medicine and technology [29,55]. For instance, the increasing presence of wearable devices and mobile health applications provides opportunities to collect richer data about a particular process, but also presents new challenges, e.g. in terms of merging all data sources [37, 55]. Even though it will require continued efforts, it is worthwhile to benefit from opportunities and tackle challenges as it will enable process mining to fully play its pivotal role to instigate evidence-based process improvement in healthcare [55].

The goal of this chapter is to introduce the reader to healthcare as an application domain for process mining. To this end, the remainder of this chapter is structured as follows. Section 2 provides a primer on healthcare processes and healthcare process data, with an emphasis on its particularities. Section 3 introduces the reader to the common use cases of process mining in healthcare from a research point of view. Section 4 discusses a case study, which illustrates the potential of process mining in the context of a specific hospital. Section 5 outlines the key open challenges that the community is confronted with when it aspires a broad uptake of process mining in healthcare. The chapter ends with a brief conclusion in Sect. 6.

2 A Primer on Healthcare Processes and Process Data

Before providing an overview of common use cases in the process mining from healthcare literature, this section sets the stage by providing an overview of healthcare organisations and healthcare processes (Sect. 2.1). Moreover, the particularities of healthcare processes and healthcare process data are introduced (Sect. 2.2).

2.1 Healthcare Organisations and Healthcare Processes

Some readers might implicitly equate healthcare to the care that patients receive in a hospital. Hospitals, either general hospitals or specialised hospitals [57], play an important role in the provision of healthcare services. As will become apparent in Sect. 3, many process mining applications are also situated within the hospital context. However, it should be noted that curative care, i.e. care focused on the treatment of diseases to increase life expectancy [88], is organised in various types of healthcare organisations [57]. For instance: long-term care facilities provide care to patients suffering from a chronic disease or patients needing longterm rehabilitation after a hospital discharge. Psychiatric care organisations, in their turn, provide therapy for patients with mental problems. Home-based care organisations, another category of healthcare organisations, deliver care services in the comfort of the patient's home [57].

Within a particular healthcare organisation, a wide variety of healthcare processes is being performed. A basic distinction between medical treatment processes and organisational processes is introduced by Lenz and Reichert [46]. *Medical treatment processes*, also commonly referred to as clinical processes, have a direct link to the patient and are connected to the therapeutic-diagnostic cycle. This implies that, in these processes, healthcare professionals takes informed decisions regarding the patient's diagnosis or therapy based on medical knowledge and the available patient-related information. *Organisational processes*, in their turn, cover all processes that support medical treatment processes by coordinating actions between different healthcare professionals and supporting staff, potentially even belonging to various departments. Examples include appointment or procedure scheduling processes, as well as logistical processes of patients or goods [46,67].

An alternative categorisation of healthcare processes is provided by Mans et al. [52]. Their classification solely takes processes that are directly related to the patients into account, but considers both medical activities as the preparation of these activities (such as booking the appointment) as being part of the same process. Against this background, Mans et al. [52] make a distinction between elective care processes and non-elective care processes. The execution of *elective care processes* can responsibly be postponed for several days or weeks. Within this subcategory, a further distinction is made between standard, routine, and non-routine care processes. For *standard care processes*, a structured treatment trajectory is available, containing information about the activities that need to be performed, as well as the timing that needs to be respected. In a routine care process, various treatment trajectories can be followed to obtain an outcome that is typically known. The latter does not hold for non-routine care processes as a physician will need to determine the next step in the treatment trajectory based on the patient's reaction on the current process step. While elective care can be postponed for several days or weeks, non-elective care processes refers to unexpected medical treatments that need to be performed promptly. Here, a distinction is made between emergency care processes, which should be executed immediately, and urgent care, which can be postponed for a limited period of time (e.g. a few days) [52].

From the previous, it follows that healthcare is a highly versatile domain, with a large variety of healthcare organisations and a mix of different processes being executed at these organisations. These processes can be fairly structured (e.g. standard care processes) or highly unstructured (e.g. non-routine care processes) [52]. The close interconnection between processes, even across different healthcare organisations, adds to the complexity of the healthcare domain. For instance: the trajectory of a patient suffering from a chronic disease might consist of surgery at a specialised hospital, several check-ups at a local general hospital, as well as multiple therapies taken at home under the supervision of a home nurse [55]. Even within a single healthcare organisation, processes are closely intertwined as, e.g., efficiently carrying out surgical processes depends on the timely execution of logistical processes, both regarding patient transportation and the material flow.

2.2 Particularities of Healthcare Processes and Process Data

To really grasp the challenging nature of healthcare as an application domain for process mining, it is important to understand the particularities of healthcare processes and healthcare process data. Munoz-Gama et al. [59] defined ten distinguishing characteristics of healthcare processes, which also impact the process data that will be recorded. While some of these characteristics might also be relevant for other sectors, their combined occurrence in the healthcare context needs to be reckoned with and will generate challenges when conducting process mining analyses. The ten key particularities of healthcare processes and healthcare process data, as defined in Munoz-Gama et al. [59], are discussed in the remainder of this subsection.

Exhibit Significant Variability. An important contributing factor to the complexity of healthcare processes is their significant variability [63,67]. Variability is caused, amongst others, by the diversity of activities that can be performed (e.g. a wide variety of examinations and treatments) in various orders, and the different characteristics of patients (e.g. they can suffer from various combinations of co-morbidities, influencing the way the process is executed) [67]. As a consequence, in many healthcare contexts, almost every case will have a unique trajectory through the process, leading to challenges within the context of, e.g., control-flow discovery [59].

Value the Infrequent Behaviour. In many domains, process mining is used to better understand the typical behaviour of a process. Hence, as infrequent behaviour would complicate, e.g., the discovered control-flow model, it is often removed in the pre-processing stage of a process mining project [15]. However, in healthcare, infrequent behaviour can be a source of valuable knowledge about the process. It might, for instance, highlight infrequent treatment paths that result in the same clinical outcome, unveiling knowledge about alternative treatment options for a particular disease [22,59]. Understanding infrequent behaviour is important as solely focusing on models representing the typical behaviour could generate blind spots, which constitute missed innovation opportunities for healthcare processes [59].

Use Guidelines and Protocols. Within the field of medicine, various clinical practice guidelines and protocols are available, which build upon evidencebased information on a certain topic [79,87]. This implies that, for clinical processes, reference processes are often available, which does not hold in many other domains [35]. This opens opportunities for process mining to, e.g., analyse the adherence to these guidelines and protocols [34,59].

Break the Glass. While clinical practice guidelines and protocols aim to achieve standardisation in clinical processes, medical doctors and healthcare professionals might need to deviate from guidelines and protocols when confronted with specific situations. For example: the discovery of specific co-morbidities of a patient might require an alternative course of action [62,72]. Another situation that might require a deviation from protocols is an unexpected surge in the number of arriving patients that should be coped with by a department [59]. The occurrence of such 'break the glass' situations will also be reflected in the data, highlighting the crucial importance to take into account context information when using process mining in healthcare to fully understand the process behaviour [59,80].

Consider Data at Multiple Abstraction Levels. In a healthcare context, data about the execution of a process can originate from various data sources, both for clinical processes and organisational processes [45,55]. These data sources will capture data at multiple levels of abstraction. Medical equipment such as surgical robots or wearable devices will often generate large volumes of very fine-grained data, which should be aggregated to retrieve meaningful patterns [59,85]. High-level data, typically recorded in administrative systems, tends to be directly interpretable, but might provide an insufficiently detailed view on the process. Hence, when performing process mining in healthcare, it might be required to integrate data from various sources, potentially bridging clinical and administrative systems, as well as different data abstraction levels [59].

Involve a Multidisciplinary Team. Healthcare processes typically have a multidisciplinary character, with healthcare professionals (physicians from various disciplines, nurses, etc.) and supporting staff with various backgrounds being involved [55,67]. Given the critical importance of expertise from the healthcare

domain, a multidisciplinary team needs to be involved during all stages of a process mining initiative, ranging from the specification of the problem to the translation of process mining insights to practical actions. This implies that attention needs to be attributed to the use of the appropriate medical terminology and customs to assure mutual understanding [59].

Focus on the Patient. When considering healthcare processes, the key role of the patient should be emphasised. Patients are, directly or indirectly, at the core of nearly all healthcare processes. Hence, when performing process mining in healthcare, specific attention should be attributed to support the provision of patient-centred care, a key care quality indicator [11]. When focusing on the patient journey, i.e. the trajectory of a patient over the course of a disease or treatment [49], it is important to note that (s)he typically receives services from various healthcare organisations (e.g. the hospital, the general practitioner, and the physiotherapist). This also causes the patient journey data to be spread over several organisations, with its associated challenges [55, 59].

Think About White-Box Approaches. Recent advances in artificial intelligence and machine learning have provided techniques to support physicians in taking complex clinical decisions. One of the biggest hurdles for the adoption of such techniques is the physician's reluctance to use systems that they do not fully understand, i.e. to use *black-box* approaches [65]. Hence, to support decisions in a healthcare context, there is a need for *white-box* approaches, enabling healthcare professionals to understand where recommendations originate from. Process mining is perceived as such a white-box approach [39]. Nevertheless, the understandability of process mining outcomes for healthcare professionals should remain a permanent point of attention [55, 59].

Generate Sensitive and Low Quality Data. Healthcare processes, especially clinical processes, generate sensitive data as it typically contain information regarding a patient's health condition, co-morbidities, ongoing treatments, etc. Consequently, ethics in general and *data privacy* in particular need to be first-class citizens when working with healthcare processes [74]. Moreover, strict regulations are typically in place regarding the use, storage and transfer of sensitive healthcare data [64]. Besides data privacy, poor data quality also characterises data collection regarding healthcare processes [54,86]. Data quality, a topic which has been discussed in Chapter 6 [18] is highly relevant in the healthcare domain, where data might suffer from various quality issues such as missing events, incorrect timestamps and imprecise timestamps [52,86]. One of the key reasons for data quality issues in healthcare is the fact that many events are recorded after a manual interaction between a healthcare professional and an information system. This might cause inaccuracies in the recorded data as some actions might not be recorded in the system, other actions might be recorded in the system well after they have been executed, etc. Data quality issues have to be handled with great care when conducting process mining in healthcare [59].

Handle Rapid Evolutions and New Paradigms. As the healthcare domain is rapidly and continuously evolving, this also holds for processes in healthcare. Changes are induced both by advances in clinical research, leading to changes in diagnostic or treatment processes [24], as well as advances in technology, e.g. the rise of remote monitoring due to the development of robust mobile health solutions [76]. New healthcare paradigms also surface, which also have an impact on healthcare processes. For instance: patient-centred care has become a core paradigm in healthcare, implying that care should attribute significant attention to the needs and preferences of the individual patient [66]. When working on process mining in healthcare, researchers and practitioners should be aware of these rapid evolutions and emerging new paradigms, as well as be able to cope with them [59].

3 Use Cases in Process Mining in Healthcare Research

Against the background of the previous section, this section aims to highlight some typical use cases for process mining in healthcare as reported in published research articles. While many of the papers that will be referenced below make important methodological contributions, the focus of the discussion in this section is mainly on how process mining techniques were applied in a particular healthcare context. To structure the outline, the six process mining types introduced in Chapter 1 [1] are used: process discovery (Sect. 3.1), conformance checking (Sect. 3.2), performance analysis (Sect. 3.3), comparative process mining (Sect. 3.4), predictive process mining (Sect. 3.5), and action-oriented process mining (Sect. 3.6). At the end of the section, some recommendations for further reading are provided (Sect. 3.7).

3.1 Process Discovery

Process discovery focuses on the discovery of a process model from an event log. As holds for process mining in general, process discovery is also, by far, the most prominent use case of process mining in healthcare [17,37]. Papers on process discovery in healthcare typically center around the discovery of the control-flow, i.e. the order of activities, from an event log [17].

When focusing on control-flow discovery, various algorithms have been used to *automatically* retrieve a visualisation of the activity order from an event log. Based on a literature review, Guzzo et al. [37] conclude that Heuristics Miner is the most commonly used algorithm, followed by Fuzzy Miner and Inductive Miner. Control-flow discovery has been applied in various healthcare contexts. For instance: Caron et al. [14] use the Heuristics Miner to retrieve a process model for the radiotherapy department within the context of gynaecologic oncology. Duma and Aringhieri [25] use both Heuristics Miner and 'Inductive Miner - Infrequent' to study the patient trajectory at the emergency department of an Italian hospital. To limit the complexity of the data, they preprocess the event log by merging consecutive events referring to the same activity in the process. Despite these pre-processing efforts, the Heuristics Miner discovers a spaghetti model, which is not understandable. The 'Inductive Miner - Infrequent', in its turn, generates a very simple, but imprecise model, meaning that the discovered model allows for a lot of behaviour that is not observed in the event log [25]. Using, amongst others, Heuristics Miner and Fuzzy Miner, Kim et al. [40] focus on the patient trajectory in an outpatient clinic in Korea. They explicitly compare the process models discovered from data to a process model that has been developed solely based on a discussion with domain experts. The process mining insights surface some important trajectories that are not included in the domain experts' model, highlighting the added value of process mining [40].

Besides automated control-flow discovery, interactive control-flow discovery also receives some attention in literature. A distinguishing characteristic of interactive control-flow discovery is that a domain expert is interactively involved while the model is being discovered from the event $\log [10]$. In this way, domain knowledge is embedded in the discovery processes, instead of being used to interpret the output of an automated algorithm. Using a case study of the patient trajectory of lung cancer patients, Benevento et al. [10] show that the interactive process discovery approach of Dixit et al. [23] generates control-flow models which are both accurate and understandable. In contrast, automated control-flow discovery algorithms might experience difficulties to generate such an accurate and understandable model. Even though the advanced algorithms discussed in Chapter 3 [7] will prove helpful, it might still be difficult to discover accurate and understandable control-flow models automatically. This can be, at least partly, explained by the fact that the order of tasks in healthcare processes often depends on highly specialised background knowledge, which is not embedded in the event log [10]. While interactive control-flow discovery received fairly little attention so far, it is highly promising for domains in which processes are highly knowledgeintensive and loosely-structured, which holds for many healthcare processes [55]. For a more extensive introduction on interactive process mining in healthcare, the reader is referred to Fernandez-Llatas [29].

A important challenge in control-flow discovery in healthcare, especially for medical treatment processes, is the great variability [59]. As many different paths through the process tend to occur, applying a control-flow discovery algorithm often results in a spaghetti model, which is very complex or even impossible to understand [51]. To handle this problem, trace clustering techniques can be used to create more homogeneous patient subgroups, which can be studied separately in an effort to reduce complexity. For instance: Mans et al. [51] use trace clustering on an event log of gynaecological oncology patients from a Dutch hospital to generate patient groups that follow a similar trajectory. Despite the potential of trace clustering, Lu et al. [48] also recognise some challenges. These include the fact that individual clusters might still contain thousands of distinct activities performed for patients, which would still be highly problematic for control-flow discovery purposes. Moreover, suppose clusters are created based on the medical condition of patients, each cluster might still contain a wide variety of patient trajectories as the same condition might be handled in a variety of ways. Against this background and with the ambition to generate clusters that are meaningful to domain experts, Lu et al. [48] develop a novel trace clustering method. Their method starts from a small sample set of patients, based on input from domain experts, to generate clusters. An evaluation of the method at a Dutch hospital highlights that the resulting control-flow models presented meaningful behavioural patterns for medical experts [48].

While the majority of control-flow discovery contributions take data from the hospital information system as a starting point, other types of input data are also occasionaly taken into consideration [37]. For example: Fernandez-Llatas et al. [31] use real-time indoor location systems data, which track the movement of patients throughout the surgery area of a Spanish hospital. Using this data, PALIA is used to discover a process model that represents the order of locations that a patient has visited [31]. Another illustration is the work of Lira et al. [47], where video recordings of a surgical procedure, i.e. the ultrasound-guided central venous catheter placement, are used as input data. These video recordings are tagged to generate an event log, which is used as an input for control-flow discovery [47].

All of the aforementioned papers focus on the discovery of control-flow models. However, as highlighted in Chapter [1] process discovery can also relate to other perspectives of the process, such as the resource perspective. For instance, Alvarez et al. [3] identify collaboration patterns between healthcare professionals within the emergency department of a hospital. The resulting process model sheds valuable insights in the interactions between physicians, nurses, medical assistants and technicians [3]. Similarly, one of the analyses conducted by Agnostinelli et al. [2] centers around the identification of interactions between different subdepartments in an Italian outpatient clinic. These examples highlight the potential of process mining to discover valuable process models in healthcare, also beyond the control-flow perspective.

3.2 Conformance Checking

As highlighted in Sect. 2.2, a multitude of clinical practice guidelines and protocols are available in the healthcare domain, which can act as reference processes [59]. Conformance checking, the topic of Chapter 5 [13] and a second common use case for process mining in healthcare, enables assessing the adherence of the real-life healthcare process (as captured by the event log) to clinical guidelines and protocols, as well as to study where reality deviates from an already existing process model [55]. For instance: Mannhardt and Blinde [50] use the public sepsis event log and aim to assess the conformance of the real-life process with two rules put forward by the sepsis guidelines at that time: (i) the time difference between the moment at which the triage document is completed and the admission of intravenous antibiotics should be less than 1 h, and (ii) the time difference between the moment at which the triage document is completed and the measurement of lactic acid should be less than three hours. Through the use of multi-perspective conformance checking, the authors conclude that the first rule is violated for 58.5% of the patients, while the second rule is only violated for 0.7% of patients. This observation constitutes a basis to look into the adherence to medical guidelines in more detail [50]. Another example is the work by Rinner et al. [68], who use alignment-based conformance checking to assess the compliance between the European guideline on melanoma treatment and an event log from an Austrian medical university. This analysis is highly relevant as the authors indicate that patients which comply to the guidelines have a significantly better prognosis than deviating patients [68]. Also focusing on clinical guidelines, Huang et al. [38] propose an approach to detect both global and local anomalies between a clinical pathway and an event log. While the former refers to patient trajectories that significantly deviate from the clinical pathway, the latter represents a deviation in a particular part of the trajectory. This approach is applied to an event log containing trajectories of unstable angina patient at a Chinese hospital [38].

While conformance checking offers great potential, Sato et al. [75] highlight the challenge that clinical guidelines and protocols are often defined at a different level of aggregation than the events in the event log. To tackle this problem and using the pre-operative phase of bariatric surgery as an illustration, the high-level activities in the reference model are explicitly mapped to the events included in the event log. Besides the potential discrepancy in terms of the level of aggregation, Bottrighi et al. [12] also highlight that clinical guidelines typically focus on patients in general, while clinical practice often requires adapting general guidelines to the specificities of individual patients and contexts. For instance: patients might have several co-morbidities and certain equipment might not be available in a particular situation. As a consequence, physicians add what is called basic medical knowledge in order to alter clinical guidelines to the specific patient and contextual characteristics. This adds a dimension to conformance checking: besides checking the adherence to the clinical guideline, the basic medical knowledge that the physician adds also needs to be taken into consideration [12].

The aforementioned examples use clinical guidelines and protocols as the reference model. While this is a common situation in the healthcare domain, it should be noted that conformance checking techniques can also generate valuable insights when the reference model originates from a different source. For instance: Kirchner et al. [41] perform conformance checking within the context of the liver transplantation process. To create the process model to compare the event log with, an interdisciplinary team consisting of physicians and modelling experts was brought together [41]. This example highlights that conformance checking is a versatile toolkit to assess whether hospital processes are performed in reality as intended according to any form of reference model.

3.3 Performance Analysis

Regarding the evaluation of healthcare process performance, various types of performance measures can be used. A basic distinction can be made between clinical, financial and operational key performance indicators. A *clinical* key performance indicator relates to a measure of the patient's medical condition, a *financial* key performance indicator reflects the financial effect of the execution of the process, and an *operational* key performance indicator represents a measure regarding the operational execution of the process. The category of operational key performance indicators can be further subdivided in *time-related* and *resource-related* key performance indicators. The former can, for example, be the waiting time of a patient or the length of stay, while the latter can relate to the bed occupancy rate or staff utilisation at a particular department [17].

Based on a systematic literature review, De Roock and Martin [17] conclude that less than half of the reviewed paper reports on a specific key performance indicator for their process mining analysis. When a key performance indicator is used, time-related key performance indicators are used the most frequently, followed by clinical key performance indicators. Financial and resource-related key performance indicators are rarely used in literature [17]. A commonly used time-related key performance indicator is the length of stay of a patient, which represents the time between the arrival of a patient and his/her departure [89].

Rojas et al. [70] use the length of stay when conducting a performance analysis of processes at the emergency department of a Chilean hospital. Based on their analysis, they identified that two key steps in the emergency department process contribute to higher length of stay values for patients. Firstly, the number of examination-treatment loops that the patient goes through, indicating the amount of time that is needed to uncover the true problem. Secondly, the need for a validation examination, which is an examination by a physician to ensure that the patient is ready to be discharged from the emergency department. In the same context and with the same key performance indicator, the length of stay at the emergency department of a hospital, Andrews et al. [5] conduct a process performance analysis at the St. Andrew's War Memorial Hospital in Australia. They conclude that a key contributor to high length of stay values is the time that elapses between the moment at which it is decided that a patient should be admitted and the moment at which the patient can actually move to the relevant ward [5].

3.4 Comparative Process Mining

Comparative process mining, e.g. the comparison of various patient groups, time periods or healthcare organisations, has also been used in the healthcare domain. With respect to the comparison of *patient groups*, Rojas and Capurro [69] study the medication use process for patients suffering from sepsis in the MIMIC-II database. To this end, three patient groups are distinguished, based on whether vasodilators, vasopressors, or systemic antibacterial antibiotics were used. Another example is Pebesma et al. [61], where three patient groups are separated to model the trajectory of cardiovascular risks for patients with type 2 diabetes: a high-risk, medium-risk and low-risk group. After modelling the evolution of the risk level for each group, the gender distribution within each group is determined, suggesting that female patients tend to be in lower risk states compared to their male counterparts. A final example is the research by Andrews et al. [6], who study the pre-hospital care process for victims of road traffic accidents. In this respect, they consider three groups: (i) persons who do not require ambulance transportation, (ii) persons who are transported to e.g. local medical practices or elderly care facilities, and (iii) persons who are transported to a hospital [6].

Other papers compare different time periods, which is another type of comparative process mining. For instance, Yoo et al. [92] use process mining to assess the impact of commissioning new buildings of a hospital, where, e.g., the cancer centre and clinical neuroscience centre have moved to the same floor and additional administrative counters have been added. To determine the impact of the move to the new building, as well as the associated new facilities that became available, the results of a process mining analysis before the move are compared to the results using an event log of a period after the move. Their findings highlight that processes run more efficiently in the new facilities, both for the cancer centre and the clinical neuroscience centre. Moreover, the consultation waiting time decreased [92]. A different example is situated within the context of an emergency department. Within that context, Stefanini et al. [77] compare the summer period to the winter period. In their comparison, they both incorporate the patients' trajectory as well as a variety of key performance indicators. One finding is that urgent patients, on average, have to wait longer before their first consultation in summer than in winter [77].

Regarding the *comparison of healthcare organisations*, a prime example is the work by Partington et al. [60]. They compare four Australian hospitals in terms of the pathway of patients who presented themselves at the emergency department and are suspected to suffer from acute coronary syndrome. The comparison focuses on the control-flow and time perspectives of the process. Regarding the time perspective, measures such as waiting times, throughput time and length of stay are taken into consideration. Various valuable insights were retrieved from the comparative analysis, e.g. some hospitals use an angiography (i.e. an X-ray of a patient's blood vessels) significantly more often than other hospitals. Moreover, significant differences in the length of stay of patients were discovered [60]. The work of Partington et al. [60] highlights the great potential of comparative process mining to compare local practices and process performance values. This can constitute a fruitful basis for mutual learning and, hence, the improvement of healthcare processes. However, it requires a culture of transparency, which has been highlighted as a challenge for process mining adoption within the broader process mining field [56].

3.5 Predictive Process Mining

While the aforementioned process mining types are backward-looking, process mining in healthcare research has also focused on forward-looking approaches, i.e. predictive process mining (see also Chapter 10 [21]). Two key research topics are data-driven prediction models and data-driven process simulation. An example of the former category, *data-driven prediction models*, is Benevento et al. [9], which focus on predicting the waiting time of patients at the emergency department. To this end, various predictor variables are taken into consideration, such as patient variables (e.g. their age or the assigned triage code), temporal variables (e.g. the hour of the day), staff-based variables (e.g. the nurses' schedules, the physicians' schedules). They also consider queue-related variables in the prediction model (e.g. the number of patients who received a triage code, but were not yet treated), which were identified in an event log. The empirical evidence suggests that adding the queue-related variables improves the performance of the waiting time prediction model. In a very different context, van der Spoel et al. [82] use a combination of data mining and process mining techniques to predict the cashflow of a Dutch hospital. In this respect, they focus on predicing the treatment trajectory based on the diagnosis and the start of the trajectory, as well as on predicting the duration of this trajectory [82].

Several papers have investigated the potential of process mining within the context of process simulation in healthcare. These efforts belong to the domain of data-driven process simulation, which refers to the extensive use of an event log during the development of a simulation model [19]. For example: Tamburis and Esposito [78] investigate how process mining could be used to support the development of a simulation model of the cataract treatment process at an ophthalmology department. Kovalchuck et al. [42], in their turn, simulate the process that patients suffering from acute coronary syndrome follow, using process mining to support the model development process. To demonstrate the developed simulation model, they focus on the effect of the availability of angiography equipment, which is important to quickly detect the presence of acute coronary syndrome. In particular, the influence of varying the number of angiography instruments on output measures such as the length of stay and the average waiting time is predicted [42]. Franck et al. [33] use a simulation-based analysis of the process of stroke patients at the emergency department. Process mining is used to determine the order of activities from an event log. Using the simulation model, various scenarios are defined in terms of the number of neurovascular intensive care unit beds required to provide patients with care according to the optimal clinical pathway.

van Hulzen et al. [84] use data-driven process simulation to explore potential future scenarios to support capacity management decisions for the radiology department of a Belgian hospital. Within the context of the construction of new facilities, which involves a centralisation of different geographically separated campuses, department management needs to provide input regarding the required number of radiological devices (X-ray, CT scanner, etc.), the size of the waiting area for ambulatory patients, and the required number of receptionists. In particular, the study centers around three key questions formulated by the department management: (i) what is the effect of the centralisation of services on the required resource capacities?, (ii) what is the impact of abolishing the need for patients to drink contrast fluid on the throughput time and required waiting area size?, and (iii) what would be the effect of an online registration system for ambulatory patients on the reception staff requirements and the size of the waiting area? To develop a simulation model to answer these questions, an event log originating from the radiology information system is intensively used. While the case study clearly demonstrates the potential of data-driven process simulation in healthcare, van Hulzen et al. [84] also highlight challenges such as data quality issues, as well as the lack of support to interactively involve domain experts during the development of a simulation model.

3.6 Action-Oriented Process Mining

As highlighted in Chapter 1 [1], action-oriented process mining focuses on translating process mining insights into actions. This is also a crucial step within the healthcare domain as only then process mining will reach its full potential as a catalyst of evidence-based process improvement [55]. Despite its great importance, research efforts focusing on the translation of process mining insights in actions are scarce in the healthcare domain. This is confirmed by the review of De Roock and Martin [17], where the need for more research on the translation of process mining outcomes to actionable process improvement ideas is indicated as one of the key recommendations for the future development of the research field.

A first step in the direction of action-oriented process mining is ensuring that process mining endeavors start from specific questions put forward by healthcare professionals [55]. Several research papers explicitly report on this matter, such as the work by van Hulzen et al. [84] on data-driven process simulation for capacity management at the radiology department. In a similar vein, Agostinelli et al. [2] explicitly devote attention to defining the questions of healthcare professionals in a process mining project in cooperation with the San Carlo di Nancy hospital. Better understanding three key processes was the central objective of the process mining analysis, including the hospitalisation process of patients. However, Agostinelli et al. [2] claimed that it was difficult to elicit specific questions from healthcare professionals because they had no background knowledge on process mining. The knowledge gap between process mining experts and domain experts is an important consideration to take into account when moving towards action-oriented process mining.

3.7 Further Reading

This section had the ambition to provide an intuitive overview of common use cases in process mining in healthcare literature. Hence, it does not constitute a full overview of all scientific contributions in the field. For a more detailed outline of the state of the art in literature, the reader is referred to one of the literature reviews on process mining in healthcare that have been published. Some reviews focus on a particular subdomain in healthcare: Kurniati et al. [43] on oncology, Kusuma et al. [44] on cardiology, Williams et al. [90] on primary care, and Farid et al. [28] on frail elderly care. Other reviews take a more generic perspective and consider process mining in healthcare as a whole: Ghasemi and Amyot [36], Rojas et al. [71], Batista and Solanas [8], Erdogan and Tarhan [27], Rule et al. [73], Dallagassa et al. [16], Guzzo et al. [37], and De Roock and Martin [17]. All review papers significantly differ in terms of the review dimensions that are taken into consideration and whether time trends are taken into consideration [17]. De Roock and Martin [17] provide an overview of the similarities and differences amongst 11 published literature reviews.

4 Case Study

The previous sections introduced healthcare processes, their particularities, and common use cases in process mining in healthcare literature. This section presents a real-life case study of conducting a process mining analysis in a hospital. The case study is situated in the *Superfluid Hospital* project conducted at the hospital of Braunschweig, led by Dr. Andreas Goepfert and Lars Anwand together with Nils Wittig. The project has the overarching ambition of ensuring that processes run smoothly within the hospital in order to improve the well-being of patients and employees, the quality of care, as well as the hospital's financial performance. To outline the case study, the project goal and IT-infrastructure is discussed (Sect. 4.1), followed by the outcomes of the process mining analysis (Sect. 4.2).

4.1 Project Goal and IT-Infrastructure

The specific goal of the *Superfluid Hospital* project is discovering medical treatment processes within the hospital. To this end, readily available process execution data and process mining has been used in order to avoid any additional documentation work for healthcare professionals. The fact that no additional data needs to be recorded could play an important role in nurturing acceptance for process mining and to stimulate its use on a continuous basis (e.g. also to track and evaluate the effect of process changes).

Hospitals typically use a variety of IT systems, implying that process execution data will also be scattered over various systems. In order to be able to analyse all relevant data centrally, the Braunschweig hospital uses data warehouse infrastructure as a starting point for process mining. This data warehouse already gathers the relevant data from various underlying information systems in the hospital. In particular, this case study uses the data warehouse infrastructure and business intelligence solution eisTIK from KMS Vertrieb und Services AG, which combines process execution data from different data sources such as the Hospital Information System, the Laboratory Information System, the Radiology Information System, etc. For the process mining analysis, an integrated version of the tool *Celonis* has been used within the data warehouse. Hence, process mining is no longer a standalone tool, which lowers the efforts for healthcare professionals to perform process mining.

4.2 Outcomes of the Process Mining Analysis

This subsection illustrates the outcomes of conducting process discovery at the case study hospital in Braunschweig. In particular, the focus will be on the medical treatment process of cardiology patients, which is a cohort of 1566 patients

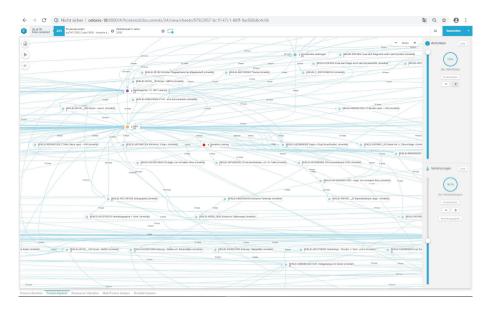


Fig. 1. Detailed view of the trajectories of patients receiving cardiology services, showing only a cut-out of the whole process.

in the data warehouse. It was the ambition of the project team, consisting of process analysts and healthcare professionals, to gain a deep understanding in the treatment of cardiology patients in order to identify areas for improvement towards the future.

Figure 1 provides an overview of the trajectories of patients receiving cardiology services, in particular a coronary angiography, containing all activities that have been conducted. As becomes apparent from the visualisation, this level of detail is unsuitable to gain insights into potential problems in the process. As a consequence, the amount of activities represented in the process model is reduced by means of filtering. Visualising only the most important activities, as shown in Fig. 2, leads to a less complex process model. The key difference between Figs. 1 and 2 is that the percentage of included activities is reduced from 100% in Fig. 1 to 53% in Fig. 2. Moreover, the number of connections between activities is also significantly reduced to about 40% in Fig. 2.

When studying Fig. 2 in more detail, it follows that particular diagnostics have already been performed for some patients before they actually go to the hospital. In particular, for 593 patients, an electrocardiogram and other check-ups ('Vorstationäre Leistungen') have already been executed before they were admitted to the hospital. Note that all results that patients bring with them will still be checked to ensure that the patient is eligible for the procedure. Patients that do not have prior check-up results generally take one of the following paths from hospital admission ('Aufnahme', blue hexagon) onwards:

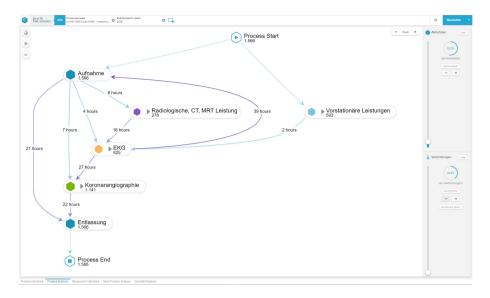


Fig. 2. Filtered view of the process for cardiology patients receiving a coronary angiography, all grouped by DRG F49G (which is a diagnosis-related grouping that is used as a billing system in Germany). (Color figure online)

- Path 1 on average 7 h to intervention: Patients following the first path directly proceed to the coronary angiography ('Koronarangiographie', green hexagon). It takes, on average, seven hours before this intervention with a coronary angiography can be performed (e.g. due to the need for a general consultation, the analysis of a blood sample, etc.). This implies that, when the patient arrives at the hospital in the morning, the intervention occurs on the same day.
- Path 2 on average 31 h to intervention: Patients following the second path receive an electrocardiogram ('EKG', yellow hexagon) on average four hours after their arrival at the hospital. When the results of the electocardiogram are available, the patient is ready for the coronary angiography. It takes, on average, 27 h before the intervention is actually carried out.
- Path 3 on average 51 h to intervention: Patients which follow the third path receive an X-ray ('Radiologische, CT, MRT Leistung', purple hexagon), on average, six hours after their arrival at the hospital. Afterwards, on average 18 h pass before the patient receives an electocardiogram ('EKG', yellow hexagon). Finally, a coronary angiography takes place, on average, another 27 h later.

Note that Fig. 2 also contains a connection between the execution of an electrocardiogram ('EKG', yellow hexagon) and hospital admission ('Aufnahme', blue hexagon). This connection represents patients which are temporarily discharged from the hospital, but return the following day to continue the process. Another interesting connection was revealed by analysing the data i.e. the direct

connection from admission ('Aufnahme', blue hexagon) to discharge (Entlassung, blue hexagon) in Fig. 2 within 21 h. This connection can be explained by the existence of a specific group of patients for whom the treatment has been recorded in a different logic. These patients have previously not been included in the internal performance measurement. This shows that process mining can also highlight relevant deviations in the documentation. In this way, important areas of action for the improvement of data quality have been identified, generating additional added value for the hospital.

As mentioned in Sect. 3, it is important that process mining insights are also translated to actions. Based on the analysis, of which some highlights have been presented above, several actions have been specified in the process, as will be exemplified here. Firstly, patients will be encouraged to bring all relevant radiological imaging and recent electrocardiogram reports with them. This will enable them to get treated much faster by following the first path described above. Secondly, measures have been taken to accelerate the second path to make sure that patients receive the intervention during their first day of hospitalisation. Due to organisational adjustments, patients now receive the ECG with higher priority. This makes it possible that, after a faster diagnosis, they often receive the actual intervention in the afternoon of the day of admission. Finally, the third path outlined above should be combined with the second path by registering patients for both the radiological and cardiological diagnostic services at the moment of admission. The relevant preliminary examinations can be carried out and evaluated over the course of a day. In this way, the procedure can take place the day after admission, provided that there are no medical reasons for not doing so.

Healthcare professionals provided positive feedback on the conducted process mining analysis, both with respect to the analysis procedure, as well as with regards to the insights that have been gathered. The conducted analysis made healthcare professionals aware of the improvement potential in their processes, which will result in shorter hospitalisations and improved care quality for patients. Especially changes that resulted in a reduction of unnecessary waiting times in the patient's trajectory are considered highly useful. While the insights and improvement actions presented in this section are based on an analysis of historical data, it should be noted that the use of the data warehouse with integrated process mining functions also enables real-time analyses. As a consequence, it is possible to create a live view of the process, which opens options to take action in the process while the process for a patient is still running.

5 Open Challenges

Section 3 and Sect. 4 demonstrate the great potential of process mining in healthcare, as well as the research that has been conducted in the research field. However, it has been reported that the uptake of process mining in healthcare, beyond case studies in a research context, is fairly limited [55]. Hence, there are still significant challenges ahead to ensure a widespread adoption of process mining in healthcare. The remainder of this section provides an overview of ten key challenges for the field, based upon the recent work by Martin et al. [55] and Munoz-Gama et al. [59].

Create a Standardised Terminology. In the healthcare domain, there is a tradition of using standardised terminologies to ensure a common understanding of concepts [26]. An illustration is the *International Classification of Diseases* (ICD), which defines about 55000 codes to label injuries, diseases, and causes of death in a standardised way [91]. In the process mining field, standardisation often focuses on the data structure level (e.g. the XES and OCEL standards), but less on the terminology level. Terms such as event, case, activity, and trace might be used in an ambiguous way based on the working definitions of individuals or research groups. This is especially troublesome when working in an interdisciplinary context as it can lead to problematic communication. Hence, there is a need to develop a standardised terminology to support process mining in healthcare, which should (i) provide a clear definition of process mining concepts in a healthcare context, and (ii) link to existing terminologies in the healthcare domain whenever possible [55].

Tackle Real-World Healthcare Problems. To support the uptake of process mining, it is important that process mining methods help to solve real-word problems of healthcare professionals. In order to capture and thoroughly understand these problems, close and ongoing interaction between the process mining community and healthcare professionals is needed. Only then, methods can be developed that actually support healthcare professionals to solve these problems [55,59]. Progress still needs to be made as, based on a systematic literature review, De Roock and Martin [17] conclude that only 12.5% of the reviewed papers reported that healthcare professionals were actively involved during the problem definition stage of a process mining project. Besides eliciting problems from healthcare professionals instead of assuming that a particular issue is relevant, it is also key to evaluate process mining methods using real-life data from an authentic healthcare context. Besides enabling the researcher to fine-tune the developed method based on the complexity of real-life data, a real-life demonstration will also build confidence among healthcare professionals in process mining's ability to tackle real-world problems [55, 59].

Deal with Low Quality Data. The healthcare domain has been shown to suffer from low quality process execution data, the key input for process mining. As applying process mining techniques to low quality data can lead to counterintuitive and even misleading results [4], data quality is an important challenge for process mining in healthcare (see also Chapter 6 [18]). Data quality issues include missing events (i.e. events that took place, but which were not registered in the system), incorrect timestamps (i.e. timestamps that do not correspond to the time at which the event actually took place), and imprecise resource information (i.e. resource information that does not refer to a specific healthcare professional) [52,54]. While approaches have recently been developed to assess the event log quality or to handle specific event log quality issues using targeted heuristics [54], data quality remains a challenge for process mining in healthcare. In this respect, it is also important that healthcare organisations are made aware of the need to improve data registration at the source in order to fully leverage the potential of process mining. Potential initiatives include raising awareness among healthcare professionals and facilitating data registration when designing user interfaces [55, 59].

Identify the Most Suitable Process Modelling Language. Within the context of control-flow discovery, process mining enables retrieving a visual representation of how a healthcare process is performed in reality. In order to effectively use a process model as a communication instrument and, hence, as a basis for process improvement, it is important to determine the most suitable process modelling language within a healthcare context. Within the business process management domain, a wide variety of process modelling languages have been developed such as BPMN, Petri nets and Declare. At the same time, modelling languages to represent clinical guidelines such as GLIF3 have been proposed in the healthcare domain. Given the plethora of available languages and as it has been shown that the modelling language impacts model understandability [32], thorough benchmarking research is required. Such research should focus on both the expressive power of the considered modelling language, as well as the understandability of the resulting control-flow model for healthcare professionals. Regarding the latter, a wide range of healthcare contexts and healthcare professionals should be taken into account. By carefully understanding the strengths and weaknesses of existing process modelling languages, both from the business process management and the healthcare domain, valuable lessons can be drawn on the visualisation of process mining outcomes in healthcare [55].

Move Beyond Control-Flow Discovery. While Sect. 3 aimed at providing a broad view on process mining in healthcare, it should be recognised that controlflow discovery remains the most dominant use case of process mining in healthcare [17,37]. While there is a clear need for control-flow discovery algorithms that are designed with the particularities of healthcare processes in mind, it is important that targeted methods are also developed for other process mining types such as conformance checking, predictive process mining or to discover insights from the time or resource perspective [59]. Moreover, as follows from Sect. 3, more research on action-oriented process mining in healthcare is needed as this is the key for process mining to actually contribute to the generation of societal value in healthcare. With respect to the various perspective of a process, analyses that span over several process perspectives, e.g. which combine the control-flow perspective with the time or resource perspective, also have the potential to generate great value for healthcare. Such multi-perspective analyses can provide healthcare professionals with rich insights, e.g. about how the control-flow of the process gives rise to particular resource behaviour [55, 59].

Look Beyond the Hospital Walls. As highlighted in Sect. 2.2, patients are at the core of healthcare processes. Patients, especially patients with a chronic disease, often have a therapeutic relationship with various healthcare organisations.

However, the great majority of the research on process mining in healthcare is still focused on what happens with patients in the context of a hospital visit or admission. Exceptions such as Fernandez-Llatas et al. [30], who focus on supporting nursing home design using process mining, are scarce. Even when a part of the patient's diagnosis and treatment process takes place in a hospital, it is important to note that a significant portion of the process might also be executed outside the hospital's walls. For instance: an oncological patient might have surgery at a specialised hospital, (s)he might have regular check-ups scheduled at a local general hospital and might receive specific treatments at home, supported by a home healthcare organisation. When process mining has the ambition to provide healthcare professionals with valuable insights in the patient journey, it will probably not be sufficient to only study the process fragment that takes place in the hospital. As process execution data will be spread over the information systems of several healthcare organisations, this will pose challenges in terms of obtaining data and connecting all data sources. Moreover, careful consideration has to be given to data privacy and security. While privacy and security are relevant for all process mining endeavours, involving several healthcare organisations will add an additional layer of complexity [55, 59].

Give Control to Healthcare Professionals. Currently, process mining initiatives in healthcare are often carried out by a multidisciplinary team, consisting of both healthcare professionals and process mining experts. Process mining experts play an important role given the technical skills which are required to prepare an event log and perform the appropriate analyses. In the long run, it should be the ambition of the process mining community to develop tools which are so intuitive that healthcare professionals can autonomously use them, instead of depending on (potentially external) process mining experts. While this is far from trivial given the high complexity of many healthcare processes, as well as due to complicating factors such as data quality issues, efforts to give control to healthcare professionals are highly valuable. A first step would, for instance, be to ensure that healthcare professionals are actively involved in the specification of analysis targets. In order to make informed judgements and clearly delineate their questions, it would be highly valuable if healthcare professionals have a minimal level of data and process literacy [2,17]. Moreover, enhanced training might also nurture a mindset in which process execution data is considered as a strategic asset that the healthcare organisation wishes to leverage to the largest extent possible. Additional efforts to gradually give control to healthcare professionals involve specific attention to elements such as the use of unambiguous terminology and the clear visualisation of outcomes when developing tools to perform process mining in healthcare [55, 59].

Integrate Process Mining Functionalities in Existing Systems. The positioning of process mining as a standalone tool constitutes a major barrier for the systematic use of process mining in healthcare practice. Nowadays, in order to use process mining, data often need to be extracted from the health information system, reformatted to the required event log structure, and imported in a process mining tool. While this is feasible for a one-off research project, this is impractical in the daily work setting of healthcare professionals. Hence, to support the use of process mining in healthcare, process mining functionalities need to be integrated in the information systems that are used by healthcare professionals. To this end, a strong partnership between the process mining community and health information system vendors needs to be established. Moreover, healthcare organisations can include the need for data-driven process analysis functions when formulating update requests to their vendors [55]. The case study presented in Sect. 4 presents a first step towards tackling this challenge as process mining functionalities were integrated with the data warehouse solution used by the hospital under consideration.

Develop Tailored Methodologies for Process Mining in Healthcare. The particularities of healthcare show the need for the development of tailored methodologies for process mining in healthcare. Such methodologies should provide specific guidelines for the various phases of a typical process mining initiative in a healthcare context, ranging from the specification of the research problem, over the composition of the event log, the execution of the analysis, to the interpretation of the final results, and the actions that will be linked to the findings. When establishing methodologies, inspiration should evidently be drawn from efforts in the broader process mining field such as the L^* -methodology [81]. and the PM²-methodology [83]. However, it is key to also take the particularities of the healthcare domain into consideration, as well as the wide variety of contexts in which process mining can be used in the domain. The presence of solid methodological support might also persuade healthcare organisations that are considering the adoption of process mining, but still have concerns regarding the rigour of a relatively young research domain, as well as regarding how the process mining effort should exactly be approached [55, 59].

Evolve in Symbiosis with Evolutions in the Healthcare Domain. As highlighted in Sect. 2.2, the healthcare domain is in constant evolution due to advances in various fields such as medicine and technology. Moreover, new paradigms such as patient-centred care give rise to new care approaches. Against this background, it will be an ongoing challenge for process mining to followup on these evolutions and to ensure that the provided support matches the expectations of healthcare organisations. To appreciate the latter statement, it is important to realise that process mining will always be a means to an end, rather than a goal in itself. Consequently, the impact of process mining in the healthcare field will depend on its ability to add value within a constantly changing context. From that perspective, process mining in healthcare should evolve in symbiosis with evolutions in the healthcare domain. While the foregoing represents a more reactive perspective, it is important to note that process mining can also actively contribute to evolutions in healthcare. For instance: process mining techniques can be used to efficiently compare various treatment processes with respect to the clinical and patient experience outcomes they generate and, hence, can contribute to shaping the clinical pathways of the future. Similarly, by providing profound insights in the usage patterns of mobile health applications, process mining can help to optimise the user-friendliness and, hence, patient satisfaction with respect to telemonitoring instruments [55, 59].

6 Conclusion

This chapter introduced a specific application domain of process mining: healthcare. Healthcare is a promising domain in which process mining can create significant societal value by helping healthcare organisations to better understand and improve their processes. Besides highlighting and illustrating the potential of various types of process mining in healthcare, the complex nature of many of its processes was also discussed. The specific characteristics of healthcare processes, such as the high level of variance and the widespread presence of guidelines and protocols, necessitate the development of dedicated process mining methods. In this respect, it is important to note that process mining in healthcare can build upon an active and committed research community, who are keen to develop novel methods that start from real-world problems experienced in healthcare. This will definitely be needed as the systematic uptake of process mining in healthcare, beyond the research context, is still fairly limited. A multitude of challenges is still ahead.

While current literature still predominantly focuses on the hospital setting, as was clearly reflected in the examples used in this chapter, it is important to also consider other types of healthcare organisations such as elderly care organisations, psychiatric care organisations and home-based care organisations. These organisations are also confronted with immense challenges and are likely to have even less resources available for advanced analytics than hospitals. Even though these other types of healthcare organisations might even be more challenging for process mining than the hospital context, e.g. because of their lower maturity in terms of data registration, they would greatly benefit from open access and user-friendly instruments from the research community to gain data-driven insights in their processes.

As a final reflection, we would like to make a message explicit that might have already become apparent while reading through this chapter: process mining in healthcare is not merely about technology and algorithms, but also about people. Actionable insights to improve healthcare processes will always emerge from the interplay between the process mining outcomes and the profound domain knowledge of healthcare professionals. Hence, it is crucial that healthcare professionals build trust in the potential of process mining and the results it generates. While healthcare professionals are a crucial actor in process mining in healthcare, another stakeholder should always remain at the center of attention: the patient. In the end, healthcare organisations, healthcare professionals, process miners and many others join forces for a single goal: to provide the best possible care to patients in a way that is sustainable in the long run. Without disregarding the numerous challenges that are still ahead, this chapter demonstrated that process mining can (and should) play an important role in achieving that goal.

References

- van der Aalst, W.M.P.: Process mining: a 360 degrees overview. In: van der Aalst, W.M.P., Carmona, J. (eds.) PMSS 2022. LNBIP, vol. 448, pp. 3–34. Springer, Cham (2022)
- Agostinelli, S., Covino, F., D'Agnese, G., De Crea, C., Leotta, F., Marrella, A.: Supporting governance in healthcare through process mining: a case study. IEEE Access 8, 186012–186025 (2020)
- 3. Alvarez, C., et al.: Discovering role interaction models in the emergency room using process mining. J. Biomed. Inform. **78**, 60–77 (2018)
- Andrews, R., Suriadi, S., Ouyang, C., Poppe, E.: Towards event log querying for data quality. In: Panetto, H., Debruyne, C., Proper, H.A., Ardagna, C.A., Roman, D., Meersman, R. (eds.) OTM 2018. LNCS, vol. 11229, pp. 116–134. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-02610-3_7
- Andrews, R., Suriadi, S., Wynn, M., ter Hofstede, A.H.M., Rothwell, S.: Improving patient flows at St. Andrew's War Memorial Hospital's emergency department through process mining. In: vom Brocke, J., Mendling, J. (eds.) Business Process Management Cases. MP, pp. 311–333. Springer, Cham (2018). https://doi.org/10. 1007/978-3-319-58307-5_17
- Andrews, R., Wynn, M.T., Vallmuur, K., Ter Hofstede, A.H.M., Bosley, E.: A comparative process mining analysis of road trauma patient pathways. Int. J. Environ. Res. Public Health 17(10), 3426 (2020)
- Augusto, A., Carmona, J., Verbeek, E.: Advanced process discovery techniques. In: van der Aalst, W.M.P., Carmona, J. (eds.) PMSS 2022. LNBIP, vol. 448, pp. 76–107. Springer, Cham (2022)
- Batista, E., Solanas, A.: Process mining in healthcare: a systematic review. In: Proceedings of the 2018 International Conference on Information, Intelligence, Systems and Applications, pp. 1–6. IEEE (2018)
- Benevento, E., Aloini, D., Squicciarini, N., Dulmin, R., Mininno, V.: Queue-based features for dynamic waiting time prediction in emergency department. Meas. Bus. Excell. 23(4), 458–471 (2019)
- Benevento, E., Dixit, P.M., Sani, M.F., Aloini, D., van der Aalst, W.M.P.: Evaluating the effectiveness of interactive process discovery in healthcare: a case study. In: Di Francescomarino, C., Dijkman, R., Zdun, U. (eds.) BPM 2019. LNBIP, vol. 362, pp. 508–519. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-37453-2_41
- Berghout, M., Van Exel, J., Leensvaart, L., Cramm, J.M.: Healthcare professionals' views on patient-centered care in hospitals. BMC Health Serv. Res. 15(1), 1–13 (2015). https://doi.org/10.1186/s12913-015-1049-z
- Bottrighi, A., Chesani, F., Mello, P., Montali, M., Montani, S., Terenziani, P.: Conformance checking of executed clinical guidelines in presence of basic medical knowledge. In: Daniel, F., Barkaoui, K., Dustdar, S. (eds.) BPM 2011. LNBIP, vol. 100, pp. 200–211. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28115-0_20
- Carmona, J., van Dongen, B., Weidlich, M.: Conformance checking: foundations, milestones and challenges. In: van der Aalst, W.M.P., Carmona, J. (eds.) PMSS 2022. LNBIP, vol. 448, pp. 155–190. Springer, Cham (2022)
- Caron, F., Vanthienen, J., Vanhaecht, K., Van Limbergen, E., De Weerdt, J., Baesens, B.: Monitoring care processes in the gynecologic oncology department. Comput. Biol. Med. 44, 88–96 (2014)

- Conforti, R., La Rosa, M., ter Hofstede, A.H.M.: Filtering out infrequent behavior from business process event logs. IEEE Trans. Knowl. Data Eng. 29(2), 300–314 (2016)
- Dallagassa, M.R., dos Santos Garcia, C., Scalabrin, E.E., Ioshii, S.O., Carvalho, D.R.: Opportunities and challenges for applying process mining in healthcare: a systematic mapping study. J. Ambient. Intell. Humaniz. Comput. 13, 165–182 (2021). https://doi.org/10.1007/s12652-021-02894-7
- De Roock, E., Martin, N.: Process mining in healthcare an updated perspective on the state of the art. J. Biomed. Inform. 127, 103995 (2022)
- De Weerdt, J., Wynn, M.T.: Foundations of process event data. In: van der Aalst, W.M.P., Carmona, J. (eds.) PMSS 2022. LNBIP, vol. 448, pp. 193–211. Springer, Cham (2022)
- Depaire, B., Martin, N.: Data-driven process simulation. In: Encyclopedia of Big Data Technologies, pp. 607–614 (2019)
- Di Ciccio, C., Marrella, A., Russo, A.: Knowledge-intensive processes: characteristics, requirements and analysis of contemporary approaches. J. Data Semant. 4(1), 29–57 (2015). https://doi.org/10.1007/s13740-014-0038-4
- Di Francescomarino, C., Ghidini, C.: Predictive process monitoring. In: van der Aalst, W.M.P., Carmona, J. (eds.) PMSS 2022. LNBIP, vol. 448, pp. 320–346. Springer, Cham (2022)
- 22. DiMatteo, M.R., Giordani, P.J., Lepper, H.S., Croghan, T.W.: Patient adherence and medical treatment outcomes a meta-analysis. Med. Care **40**(9), 794–811 (2002)
- Dixit, P.M., Verbeek, H.M.W., Buijs, J.C.A.M., van der Aalst, W.M.P.: Interactive data-driven process model construction. In: Trujillo, J.C., et al. (eds.) ER 2018. LNCS, vol. 11157, pp. 251–265. Springer, Cham (2018). https://doi.org/10.1007/ 978-3-030-00847-5_19
- Djulbegovic, B., Guyatt, G.H.: Progress in evidence-based medicine: a quarter century on. The Lancet **390**(10092), 415–423 (2017)
- Duma, D., Aringhieri, R.: Mining the patient flow through an emergency department to deal with overcrowding. In: Cappanera, P., Li, J., Matta, A., Sahin, E., Vandaele, N.J., Visintin, F. (eds.) ICHCSE 2017. SPMS, vol. 210, pp. 49–59. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66146-9_5
- 26. Engelhorn, M.: Semantics and big data semantics methods for data processing and searching large amounts of data. In: Langkafel, P. (ed.) Big Data in Medical Science and Healthcare Management, pp. 177–196. Walter de Gruyter, Berlin (2016)
- 27. Erdogan, T.G., Tarhan, A.: Systematic mapping of process mining studies in healthcare. IEEE Access 6, 24543–24567 (2018)
- Farid, N.F., De Kamps, M., Johnson, O.A.: Process mining in frail elderly care: a literature review. In: Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies-Volume 5: HEALTHINF, vol. 5, pp. 332–339. SciTePress, Science and Technology Publications (2019)
- Fernandez-Llatas, C.: Interactive Process Mining in Healthcare. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-53993-1
- 30. Fernández-Llatas, C., et al.: Behaviour patterns detection for persuasive design in nursing homes to help dementia patients. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6413–6417. IEEE (2011)
- Fernandez-Llatas, C., Lizondo, A., Monton, E., Benedi, J.-M., Traver, V.: Process mining methodology for health process tracking using real-time indoor location systems. Sensors 15(12), 29821–29840 (2015)

- 32. Figl, K.: Comprehension of procedural visual business process models. Bus. Inf. Syst. Eng. 59(1), 41–67 (2017). https://doi.org/10.1007/s12599-016-0460-2
- 33. Franck, T., Bercelli, P., Aloui, S., Augusto, V.: A generic framework to analyze and improve patient pathways within a healthcare network using process mining and discrete-event simulation. In: Proceedings of the 2020 Winter Simulation Conference, pp. 968–979. IEEE (2020)
- 34. Gatta, R., et al.: Clinical guidelines: a crossroad of many research areas. Challenges and opportunities in process mining for healthcare. In: Di Francescomarino, C., Dijkman, R., Zdun, U. (eds.) BPM 2019. LNBIP, vol. 362, pp. 545–556. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-37453-2_44
- Gatta, R., et al.: What role can process mining play in recurrent clinical guidelines issues? A position paper. Int. J. Environ. Res. Public Health 17(18), 6616 (2020)
- Ghasemi, M., Amyot, D.: Process mining in healthcare: a systematised literature review. Int. J. Electron. Healthc. 9(1), 60–88 (2016)
- Guzzo, A., Rullo, A., Vocaturo, E.: Process mining applications in the healthcare domain: a comprehensive review. Wiley Interdisc. Rev. Data Min. Knowl. Discov. 12(2), e1442 (2022)
- Huang, Z., Dong, W., Ji, L., Yin, L., Duan, H.: On local anomaly detection and analysis for clinical pathways. Artif. Intell. Med. 65(3), 167–177 (2015)
- Jans, M., Soffer, P., Jouck, T.: Building a valuable event log for process mining: an experimental exploration of a guided process. Enterp. Inf. Syst. 13(5), 601–630 (2019)
- Kim, E., et al.: Discovery of outpatient care process of a tertiary university hospital using process mining. Healthc. Inform. Res. 19(1), 42–49 (2013)
- Kirchner, K., Herzberg, N., Rogge-Solti, A., Weske, M.: Embedding conformance checking in a process intelligence system in hospital environments. In: Lenz, R., Miksch, S., Peleg, M., Reichert, M., Riaño, D., ten Teije, A. (eds.) KR4HC/ProHealth - 2012. LNCS (LNAI), vol. 7738, pp. 126–139. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36438-9_9
- Kovalchuk, S.V., Funkner, A.A., Metsker, O.G., Yakovlev, A.N.: Simulation of patient flow in multiple healthcare units using process and data mining techniques for model identification. J. Biomed. Inform. 82, 128–142 (2018)
- Kurniati, A.P., Johnson, O., Hogg, D., Hall, G.: Process mining in oncology: a literature review. In: Proceedings of the 2016 International Conference on Information Communication and Management, pp. 291–297. IEEE (2016)
- 44. Kusuma, G.P., Hall, M., Gale, C.P., Johnson, O.A.: Process mining in cardiology: a literature review. Int. J. Biosci. Biochem. Bioinform. 8(4), 226–236 (2018)
- Lenz, R., Peleg, M., Reichert, M.: Healthcare process support: achievements, challenges, current research. Int. J. Knowl.-Based Organ. 2(4) (2012)
- Lenz, R., Reichert, M.: IT support for healthcare processes premises, challenges, perspectives. Data Knowl. Eng. 61(1), 39–58 (2007)
- 47. Lira, R., et al.: Process-oriented feedback through process mining for surgical procedures in medical training: the ultrasound-guided central venous catheter placement case. Int. J. Environ. Res. Public Health 16(11), 2019 (1877)
- Lu, X., Tabatabaei, S.A., Hoogendoorn, M., Reijers, H.A.: Trace clustering on very large event data in healthcare using frequent sequence patterns. In: Hildebrandt, T., van Dongen, B.F., Röglinger, M., Mendling, J. (eds.) BPM 2019. LNCS, vol. 11675, pp. 198–215. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-26619-6_14
- Manchaiah, V.K.C., Stephens, D., Meredith, R.: The patient journey of adults with hearing impairment: the patients' views. Clin. Otolaryngol. 36(3), 227–234 (2011)

- 50. Mannhardt, F., Blinde, D.: Analyzing the trajectories of patients with sepsis using process mining. In: CEUR Workshop Proceedings, vol. 1859, pp. 72–80 (2017)
- Mans, R.S., Schonenberg, M.H., Song, M., van der Aalst, W.M.P., Bakker, P.J.M.: Process mining in healthcare. In: Proceedings of the 2008 International Conference on Health Informatics, pp. 118–125 (2008)
- Mans, R.S., van der Aalst, W.M.P., Vanwersch, R.J.B.: Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-319-16071-9
- Mans, R.S., van der Aalst, W.M.P., Russell, N.C., Bakker, P.J.M., Moleman, A.J.: Process-aware information system development for the healthcare domain - consistency, reliability, and effectiveness. In: Rinderle-Ma, S., Sadiq, S., Leymann, F. (eds.) BPM 2009. LNBIP, vol. 43, pp. 635–646. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12186-9_61
- Martin, N.: Data quality in process mining. In: Fernandez-Llatas, C. (ed.) Interactive Process Mining in Healthcare. HI, pp. 53–79. Springer, Cham (2021). https:// doi.org/10.1007/978-3-030-53993-1_5
- 55. Martin, N., et al.: Recommendations for enhancing the usability and understandability of process mining in healthcare. Artif. Intell. Med. **109**, 101962 (2020)
- 56. Martin, N., et al.: Opportunities and challenges for process mining in organizations: results of a Delphi study. Bus. Inf. Syst. Eng. 63(5), 511–527 (2021). https://doi. org/10.1007/s12599-021-00720-0
- 57. Maurer, F.A., Smith, C.M.: Community/Public Health Nursing Practice: Health for Families and Populations. Elsevier Saunders, St. Louis (2013)
- Mertens, S., Gailly, F., Poels, G.: Towards a decision-aware declarative process modeling language for knowledge-intensive processes. Expert Syst. Appl. 87, 316– 334 (2017)
- Munoz-Gama, J., et al.: Process mining for healthcare: characteristics and challenges. J. Biomed. Inform. 127, 103994 (2022)
- Partington, A., Wynn, M., Suriadi, S., Ouyang, C., Karnon, J.: Process mining for clinical processes: a comparative analysis of four Australian hospitals. ACM Trans. Manag. Inf. Syst. 5(4), 1–18 (2015)
- Pebesma, J., et al.: Clustering cardiovascular risk trajectories of patients with type 2 diabetes using process mining. In: Proceedings of the 2019 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 341–344. IEEE (2019)
- Peleg, M.: Computer-interpretable clinical guidelines: a methodological review. J. Biomed. Inform. 46(4), 744–763 (2013)
- Pereira Detro, S., Santos, E.A.P., Panetto, H., De Loures, E., Lezoche, M., Cabral Moro Barra, C.: Applying process mining and semantic reasoning for process model customisation in healthcare. Enterp. Inf. Syst. 14(7), 983–1009 (2020)
- Pika, A., Wynn, M.T., Budiono, S., ter Hofstede, A.H.M., van der Aalst, W.M.P., Reijers, H.A.: Privacy-preserving process mining in healthcare. Int. J. Environ. Res. Public Health 17(5), 1612 (2020)
- Poon, A.I.F., Sung, J.J.Y.: Opening the black box of AI-medicine. J. Gastroenterol. Hepatol. 36(3), 581–584 (2021)
- Rathert, C., Wyrwich, M.D., Boren, S.A.: Patient-centered care and outcomes: a systematic review of the literature. Med. Care Res. Rev. 70(4), 351–379 (2013)
- 67. Rebuge, Á., Ferreira, D.R.: Business process analysis in healthcare environments: a methodology based on process mining. Inf. Syst. **37**(2), 99–116 (2012)

- Rinner, C., Helm, E., Dunkl, R., Kittler, H., Rinderle-Ma, S.: An application of process mining in the context of melanoma surveillance using time boxing. In: Daniel, F., Sheng, Q.Z., Motahari, H. (eds.) BPM 2018. LNBIP, vol. 342, pp. 175– 186. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11641-5_14
- Rojas, E., Capurro, D.: Characterization of drug use patterns using process mining and temporal abstraction digital phenotyping. In: Daniel, F., Sheng, Q.Z., Motahari, H. (eds.) BPM 2018. LNBIP, vol. 342, pp. 187–198. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11641-5_15
- Rojas, E., Cifuentes, A., Burattin, A., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Analysis of emergency room episodes duration through process mining. In: Daniel, F., Sheng, Q.Z., Motahari, H. (eds.) BPM 2018. LNBIP, vol. 342, pp. 251– 263. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11641-5_20
- Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: a literature review. J. Biomed. Inform. 61, 224–236 (2016)
- Roulin, D., Muradbegovic, M., Addor, V., Blanc, C., Demartines, N., Hübner, M.: Enhanced recovery after elective colorectal surgery-reasons for non-compliance with the protocol. Dig. Surg. 34(3), 220–226 (2017)
- Rule, A., Chiang, M.F., Hribar, M.R.: Using electronic health record audit logs to study clinical activity: a systematic review of aims, measures, and methods. J. Am. Med. Inform. Assoc. 27(3), 480–490 (2020)
- Safran, C., et al.: 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), 1–9 (2007)
- 75. Sato, D.M.V., de Freitas, S.C., Dallagassa, M.R., Scalabrin, E.E., Portela, E.A.P., Carvalho, D.R.: Conformance checking with different levels of granularity: a case study on bariatric surgery. In: Proceedings of the 2020 International Congress on Image and Signal Processing, Biomedical Engineering and Informatics, pp. 820– 826. IEEE (2020)
- Silva, B.M.C., Rodrigues, J.J.P.C., de la Torre Díez, I., López-Coronado, M., Saleem, K.: Mobile-health: a review of current state in 2015. J. Biomed. Inform. 56, 265–272 (2015)
- Stefanini, A., Aloini, D., Benevento, E., Dulmin, R., Mininno, V.: Performance analysis in emergency departments: a data-driven approach. Meas. Bus. Excell. 22(2), 130–145 (2018)
- Tamburis, O., Esposito, C.: Process mining as support to simulation modeling: a hospital-based case study. Simul. Model. Pract. Theory 104, 102149 (2020)
- Ten Teije, A., et al.: Improving medical protocols by formal methods. Artif. Intell. Med. 36(3), 193–209 (2006)
- van Andel, V., Beerepoot, I., Lu, X., van de Weerd, I., Reijers, H.A.: DEUCE: a methodology for detecting unauthorized access of electronic health records using process mining. In: Proceedings of the European Conference on Information Systems, p. 1340 (2021)
- van der Aalst, W.M.P.: Process Mining: Data Science in Action. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49851-4
- van der Spoel, S., van Keulen, M., Amrit, C.: Process prediction in noisy data sets: a case study in a Dutch hospital. In: Cudre-Mauroux, P., Ceravolo, P., Gašević, D. (eds.) SIMPDA 2012. LNBIP, vol. 162, pp. 60–83. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40919-6_4
- van Eck, M.L., Lu, X., Leemans, S.J.J., van der Aalst, W.M.P.: PM²: a process mining project methodology. In: Zdravkovic, J., Kirikova, M., Johannesson, P.

(eds.) CAiSE 2015. LNCS, vol. 9097, pp. 297–313. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-19069-3_19

- van Hulzen, G., Martin, N., Depaire, B., Souverijns, G.: Supporting capacity management decisions in healthcare using data-driven process simulation. J. Biomed. Inform. 129, 104060 (2022)
- van Zelst, S.J., Mannhardt, F., de Leoni, M., Koschmider, A.: Event abstraction in process mining: literature review and taxonomy. Granular Comput. 6(3), 719–736 (2020). https://doi.org/10.1007/s41066-020-00226-2
- Vanbrabant, L., Martin, N., Ramaekers, K., Braekers, K.: Quality of input data in emergency department simulations: framework and assessment techniques. Simul. Model. Pract. Theory **91**, 83–101 (2019)
- Wang, D., et al.: Representation primitives, process models and patient data in computer-interpretable clinical practice guidelines: a literature review of guideline representation models. Int. J. Med. Inform. 68(1-3), 59-70 (2002)
- Wang, F.: The roles of preventive and curative health care in economic development. PLoS ONE 13(11), e0206808 (2018)
- Wiler, J.L., Welch, S., Pines, J., Schuur, J., Jouriles, N., Stone-Griffith, S.: Emergency department performance measures updates: proceedings of the 2014 emergency department benchmarking alliance consensus summit. Acad. Emerg. Med. 22(5), 542–553 (2015)
- Williams, R., Rojas, E., Peek, N., Johnson, O.A.: Process mining in primary care: a literature review. Stud. Health Technol. Inform. 247, 376–380 (2018)
- World Health Organization: WHO releases new international classification of diseases (ICD 11) (2018). https://www.who.int/news-room/detail/18-06-2018-whoreleases-new-international-classification-of-diseases-(icd-11). Accessed 08 Apr 2022
- Yoo, S., et al.: Assessment of hospital processes using a process mining technique: outpatient process analysis at a tertiary hospital. Int. J. Med. Inform. 88, 34–43 (2016)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

