

An Application of Process Mining in the Context of Melanoma Surveillance using Time Boxing

Christoph Rinner¹[0000-0002-3847-5664], Emmanuel Helm²[0000-0002-1323-3511],
Reinhold Dunkl³, Harald Kittler⁴[0000-0002-0051-8016], and Stefanie
Rinderle-Ma³[0000-0001-5656-6108]

¹ Medical University of Vienna, Center for Medical Statistics, Informatics, and Intelligent Systems (CeMSIIS), Spitalgasse 23, Vienna, Austria

² University of Applied Sciences Upper Austria, Research Department of Advanced Information Systems and Technology, Softwarepark 13, Hagenberg, Austria

³ University of Vienna, Faculty of Computer Science, Währinger Strasse 29, Vienna, Austria

⁴ Medical University of Vienna, Department of Dermatology, Währinger Gürtel 18-20, Vienna, Austria

Abstract. Background: Process mining is a relatively new discipline that helps to discover and analyze actual process executions based on log data. In this paper we apply conformance checking techniques to the process of surveillance of melanoma patients. This process consists of recurring events with time constraints between the events. Objectives: The goal of this work is to show how existing clinical data collected during melanoma surveillance can be prepared and pre-processed to be reused for process mining. Methods: We describe an approach based on time boxing to create process models from medical guidelines and the corresponding event logs from clinical data of patient visits. Results: Event logs were extracted for 1,023 patients starting melanoma surveillance at the Department of Dermatology at the Medical University of Vienna between January 2010 and June 2017. Conformance checking techniques available in the ProM framework were applied. Conclusions: The presented time boxing enables the direct use of existing process mining frameworks like ProM to perform process-oriented analysis also with respect to time constraints between events.

Keywords: health care processes · process mining · electronic health records · medical guidelines.

1 Introduction

1.1 Motivation

The availability of electronic data in the early 1980s changed the business models of many companies as well as the medical domain. The proliferation of electronic health data from a variety of sources including medical treatment records, administrative health data and public health information systems supports the

integration of all available evidence to feed back into medical research and practice. These health data collected during routine care can be used for secondary purposes such as identifying trends, predicting outcomes, influencing patient care, drug development and therapy choices [23]. By analyzing these data it is possible to gain new insights which can be used for the optimization of health care processes, i.e., the insights might enable the transformation of health care processes instead of simply monitoring them.

Skin malignancies are recognized as a major and global health problem. Accounting for about 5% of all skin cancer cases melanoma is the most dangerous form of skin malignancy and causes about 90% of skin cancer mortalities [9]. Incidence rates in many European countries are actually ranging between 12 and 15 cases per 100.000 inhabitants. Currently, the increasing rates are levelling off in some countries. In contrast, however, for distinct subpopulations such as elderly men the rates are still increasing [11]. Early detection of melanoma is of utmost importance and is leading to a favourable prognosis. Since melanoma may appear years after the excision of the primary tumor, patients with melanoma are monitored closely, usually following a predefined protocol, to allow timely detection of recurrent disease [7].

1.2 Problem Statement

In order to improve the surveillance of patients with melanoma traditional studies depended on manual data acquisition. The goal of this work is to show how existing data from routine care can be combined from different sources and reused for process mining to automatically detect processes [1] and compare them to medical guidelines using conformance checking [14, 22].

Process mining is a relatively new discipline that helps to discover and analyze actual process executions based on log data. Log data stores events that are produced during process execution, for example, the execution of a process activity excision. Process mining techniques are particularly promising to be applied in the healthcare domain facing challenges to a) learn the process of interest; b) understand deviations, c) analyze bottlenecks, and d) monitor organizational behaviour [16].

1.3 Research Questions

This paper focuses on challenges a) and b) and aims at learning about the applicability and results of process mining and conformance checking for the treatment and surveillance of melanomas. Previous studies have indicated the potential of process mining in this area [3, 6]. In these two studies, however, only a maximum of 10 process instances were analyzed. Moreover, several data challenges were pointed out, specifically that the time granularity of the logged data was too coarse [3]. This problem is also referred to by [16] as imprecise data. Hence, this paper addresses the following research questions:

1. How can existing clinical data be reused for the application of process mining?

2. How can data of recurring events with time constraints that span a long period of time be prepared to apply process mining?
3. How can we apply process mining to check guideline compliance?
4. What can we learn from process mining in the context of surveillance of melanoma patients?

1.4 Contribution

This study provides conceptual extensions towards data preparation for imprecise log data. In particular, we describe a method for data preparation using a specific naming convention to model the time aspects used in medical guidelines (e.g. Check up in 6 months). The methods were tested using follow-up guidelines for melanoma patients and anonymous patient data from the Department of Dermatology at the Medical University Vienna (DDMUV).

In Section 2 we give a brief overview of the process mining discipline and the positioning of this paper in the field of process mining in healthcare. In Section 3 we describe the data preparation and time boxing steps as well as the conformance checking applied in this paper. In Section 4 the study population at the DDMUV is described and time boxing and conformance checking are applied to melanoma surveillance data. In Section 5 the generalized applicability of our methods and the medical implications are discussed.

2 Related Work

2.1 Process-oriented analysis

Process mining [1] offers techniques for different analysis tasks such as *discovery* of process models, *conformance checking* between process execution logs and process models or guidelines, and *enhancement* of existing models using information about the recorded reality in process execution logs (i.e. event logs). Event logs store events that are produced by the execution of the process tasks during runtime. Existing standardized formats for event logs are MXML and XES [25]. In this article we mainly focus on the task of *conformance checking*.

There are different techniques and algorithms to measure conformance (e.g. [22, 2]). The degree of conformance is measured in four orthogonal dimensions. (1) Fitness is the most agreed upon measure to determine if the model reflects the recorded behaviour in the event log. Approaches to measure fitness are listed and evaluated by [4]. (2) Precision aims to identify overly general models and penalize unwanted behaviour. A recent study, however, shows that existing precision metrics do not provide the desired properties to reliably recognize under-fitting [24]. (3) Generalization measures the degree of overfitting, i.e. if the model only allows for what has actually been observed. (4) Structural appropriateness, derived from the simplicity quality dimension described in [1], aims to find an easy to understand and not overly complex model.

2.2 Positioning in the field

Rojas et al. provide a recent survey on existing literature and case studies on process mining in healthcare, collecting 74 publications in this area [20]. Following their terminology, our paper analyzes *organizational* processes based on data from a *clinical support system* for *oncology* in *Austria*. It poses a *specific question* (i.e. guideline compliance) and utilizes the *conformance* perspective. This paper presents a *semi-automated* implementation strategy, providing a novel data preparation approach facilitating the use of the *tool* ProM. The analysis follows the *basic* approach, using the ProM plugin Multi-perspective Process Explorer (MPE) [14].

3 Methods

The process of surveillance of melanoma patients starts with the detection and the excision of the primary tumor (i.e. the baseline visit). Melanoma patients are staged according to the American Joint Committee on Cancer (AJCC) staging system (i.e. stage I to IV). After excision of the primary tumor patients start a 10 year surveillance period. Depending on AJCC staging, follow-up visits have different surveillance intervals and include different types of examinations (i.e. clinical examination, analyzing tumor markers, lymph node sonography, computed tomography of the abdomen, PET-CT etc.). In AJCC stage I for example the interval of the follow-up visits is 6 months in the first 5 years and one year between the 5th and the 10th year. In the other AJCC stages intervals of 3 months in the first five years and 6 month between the 5th and the 10th year are scheduled. The higher the AJCC stage the more often examinations are performed as part of a follow-up program. During surveillance, the AJCC stages are re-evaluated and patients can be assigned a higher AJCC stage and start the corresponding follow-up surveillance from the beginning. In this paper we refer to this upgrading as stage change. Since the observation period of this study only covers 7 years (January 2010 to June 2017) patients are still compliant to the guideline if they have not missed the next-to-last or last follow-up visit before the end of the study (i.e. June 2017). Patients are considered lost to follow-up when the surveillance is terminated prematurely at the DDMUV (e.g. patient changed clinic). The events occurring during melanoma surveillance are depicted in Figure 1.

We took the clinical data needed to perform the process mining from a local melanoma registry stored in the Research and Analysis (RDA) Platform of the Medical University of Vienna [5]. Additional information about the transfer of patients between different clinics within the hospital, laboratory results as well as treatment information is obtained from the local Hospital Information System (HIS). Since the melanoma registry is maintained manually the information from the HIS is used to detect additional follow-up visits re-using information from routine care. The event logs used in this study are created using the JAVA programming language, a JDBC driver to access the data in the Oracle Database and the OpenXES library to create event logs in the MXML format.

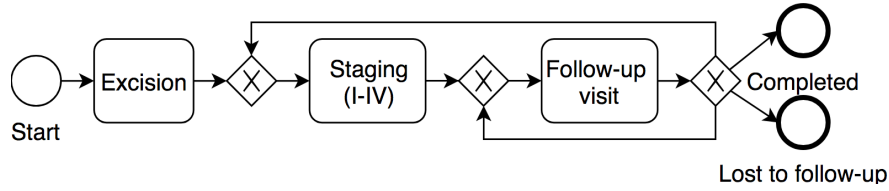


Fig. 1. A BPMN representation of the process of melanoma surveillance. The start event of the surveillance is the excision of the melanoma followed by the AJCC stage classification and the follow-up visits. Depending on the AJCC stage the number of follow-up visits can vary and the AJCC stage can be re-assessed after each follow-up visit. A patient can be lost to follow-up or complete the surveillance successfully.

Conformance checking was performed in the ProM framework. The study was approved by the ethics committee of the Medical University of Vienna (EK Nr.: 1297/2014).

3.1 Data preparation and time boxing

According to the guideline, depending on the AJCC stage of the patient, the follow-up treatment takes place at certain time intervals (i.e. every three, six or twelve months) in a repeated fashion for ten years. The existing process execution logs record the same event for each occurrence (e.g. *follow-up visit* for each follow-up visit), so the process mining algorithms are not able to distinguish between these events depending on the fixed time period specified in the guideline (e.g. second follow-up visit after one year).

Using the simplified process model depicted in Figure 1, it is not possible to distinguish the different follow-up visits automatically and as a consequence conformance to the guideline cannot be checked using current process mining algorithms. To overcome this problem we propose a naming convention based on *time boxing* for recurring events commonly described in medical guidelines. During the time boxing, each activity (e.g. each follow-up visit) is allocated (i.e. aligned) to a predefined fixed time period it matches in, called a time box. Each time box corresponds to an event in the medical guideline and the events in each time box are named according to the name of the time box (e.g. *IF_01_1Q* corresponds to an AJCC stage I follow-up visit in the first quarter of the first year). In Figure 2 the process model with applied time boxing corresponding to the melanoma surveillance guideline used at the DDMUV is shown.

The event log follows the same naming convention as the process model. To generate the event log, all follow-up visits are assigned to the corresponding (i.e. temporally closest) time boxes. All follow-up visits in one time box are merged and represented as one. In order to analyze over-compliance, multiple events could be assigned to the same time box (without merging) and the resulting self-loops considered during the analysis.

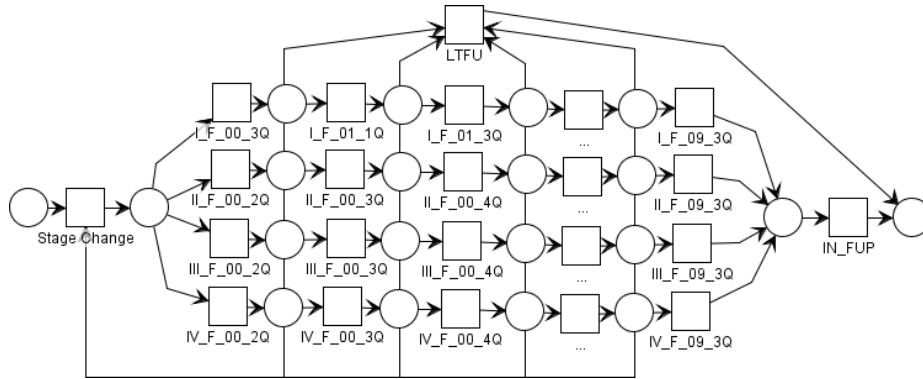


Fig. 2. A simplified petri net model with applied time boxing corresponding to the guideline used at the DDMUV. For each AJCC stage (i.e. I, II, III and IV) the follow-up visits after three, six or twelve month (i.e. 2Q is 2nd quarter, 3Q is 3rd quarter 4Q is 4th quarter) for ten years are shown. After each follow-up visit a patient can proceed to any later follow-up visit, have a state change, be lost to follow-up (LTFU) or complete the surveillance (i.e. IN_FUP meaning *still in follow-up*).

3.2 Conformance checking with ProM

The conformance checking was done using the process mining framework ProM in version 6.6 and the respective plug-in *Multi-perspective Process Explorer* (MPE) [14]. It allows for fitness and precision calculation and provides different views on the data, including (1) a model view, depicting the base model petri net, (2) a trace view, making it possible to investigate individual traces and (3) a chart view, showing the distribution of attribute values in the log for certain parts of the model.

The MPE is an advanced tool that integrates state of the art algorithms described in [15], that are also able to integrate different perspectives i.e. data, resource and time. The configuration for penalties on log and event moves was adapted to the specific use case. A valid configuration for the alignment parameters (penalties for moves on the log/model) had to be identified. Due to the pre-processing there are no wrong events (events present in the log but not in the model) save for the LTFU (i.e. Lost to follow-up) event, so the alignment algorithm must always identify the missing events (events in the model that are missing in the log). The parameters are described in the results section.

4 Results

The DDMUV is a tertiary referral centre that offers a long-term surveillance program for melanoma patients based on the European guideline on melanoma treatment [9]. An example for the follow-up sub process in the European guideline on melanoma treatment modelled in the Business Process Modeling and

Notation (BPMN) can be found in [3]. The melanoma registry at the DDMUV contains data of baseline and follow-up visits of melanoma patients. Excisions are documented way back to the early 1990s, a continuous documentation of the follow-up visits started 2010. In 2017, the melanoma registry covered about 2,200 patients. In this study we included all 1023 patients (43% females, mean age 59 ± 17.5 years) with baseline visit (i.e. excisions) after January 2010 and at least one follow-up visit. Besides the demographic data, different characteristics of the identified melanoma are documented. For the baseline visit this includes among others, (1) melanoma subtype (superficial spreading melanoma, nodular melanoma, lentigo maligna melanoma, acral lentiginous melanoma and others), (2) anatomic site (i.e. abdomen, hand, foot, head etc.), (3) depth of invasion, (4) date of surgery for the primary tumor and (5) staging information. More than one primary tumor can be documented. Only melanoma staging is used for conformance checking. We extracted five different event logs from our real world data, one including all patients (i.e. I-IV), and four for each AJCC stage separately (i.e. I, II, III, IV) based on the highest AJCC stage of the patient. If a patient initially started with AJCC stage I and then moved to AJCC stage II the patient is represented in the AJCC II log file. Table 1 lists the number of patients in each log.

Table 1. Number of patients in log files split by maximal AJCC stage and separated by LTFU (lost to follow-up) and IN_FUP (in follow-up). The distribution of the outcome indicators LTFU and IN_FUP show significant differences between the AJCC stages.

		I-IV	I	II	III	IV
Female	LTFU	286	146	95	20	25
	IN_FUP	153	45	50	17	41
	Total	439	191	145	37	66
Male	LTFU	379	167	124	42	46
	IN_FUP	205	43	70	33	59
	Total	584	210	194	75	105
Total	LTFU	665	313	219	62	71
	IN_FUP	358	88	120	50	100
	Total	1023	401	339	112	171

The number of patients per AJCC stage decreases with higher AJCC stage, which corresponds to the fact that most melanomas in Austria are diagnosed in early stages [10]. Most patients ($n=401$) were in AJCC stage I. This group also had the highest number of patients lost to follow-up ($n=313$, 78%). The ratio of patients IN_FUP (i.e. in follow-up) was the highest in AJCC stage IV with 58% ($n=100$). There is no difference between proportion of individuals lost to follow-up (LTFU) between men and women. Men were generally older than women and there was no significant difference between the LTFU and IN_FUP in respect to the age. Patients in lower AJCC stages were generally younger (I:

mean age 57 ± 17 years; II: mean age 59 ± 18 years; III: mean age 60 ± 18 years; IV: mean age 63 ± 16 years)

4.1 Conformance checking of melanoma surveillance

To check the conformance of our guideline models in regard to the recorded event logs, we replayed the logs on the models using the MPE. For the alignment, the default costs of the MPE for missing events in the log (value: 2) and missing activities in the model (value: 3) leads to undesired behaviour. The alignment algorithm identifies follow-up visits after a long period of time as wrong events. When the penalty for a sequence of missing events exceeds the penalty for a wrong event, the alignment algorithm will declare the current event wrong. In order to ensure a correct alignment, the maximum number of skipped follow-up visits in all traces was identified and the penalties adopted respectively. Since the maximum number of consecutive skipped events for one trace is 19 in our data, we chose a penalty of 1 for missing events and 20 for wrong events. For the LTFU event we reduce the wrong event penalty to 0, thus only penalizing the missing IN_FUP event at the end and not overvalue the outcome indicator for the fitness calculation. The results in the form of fitness and precision indicators can be seen in Table 2.

Table 2. Using the MPE, for each stage log and the combined log (I-IV) the average fitness and precision were calculated in regard to the respective guideline models.

AJCC stage	No of patients	Avg. fitness % (min - max)	Avg. precision % (#observed / #possible behaviour)
I	401	98,6% (91% - 100%)	75,1%
II	339	98,0% (82% - 100%)	71,4%
III	112	98,2% (85% - 100%)	65,0%
IV	171	98,7% (88% - 100%)	63,1%
I-IV	1.023	98,4% (53% - 100%)	87,0%

Our measurements show that the guideline models have an overall comparable and good fitness value, i.e. the model generally explains the behaviour seen in the log. This originates from three facts. (1) The renaming and clustering of activities was done based on the terminology that was also used for the guideline model. (2) The time boxing method presented in 3.1 leads to an ordered sequence of events, where loops and duplicates cannot occur. (3) The only *wrong* events (i.e. events present only in the log, not in the model) are the LTFU events.

The precision of the model for stage I is 75.1% and declines to 63.1% for stage IV. The ratio between observed and possible behaviour indicates underfitting for low values. The explanation for the generally lower precision values is that the guideline models include the whole time period of ten years of follow-up visits, while the event logs only cover a maximum of seven and a half years. Thus, modelled events like I.F_08_1Q (i.e. stage I, eighth year, first quarter) will

5 Discussion and lessons learnt

5.1 Reuse of clinical data for process mining

We reused existing patient data available from a local EHR system in the context of melanoma surveillance. In combination with the local melanoma registry additional follow-up visits and laboratory data to the event log were identified. Creating the log file using a procedural programming approach allowed us to add pre-processing steps. For example we tagged patients that successfully terminate the process (i.e. still in the surveillance program (IN_FUP)) during the creation of the log file based on the time they did not show up before the end (i.e. time boxes missed). Beside the melanoma registry, more than 70 other registries are documented in the RDA platform. In a current master thesis a mapping of the melanoma registry data from the RDA data model [19] to the i2b2 star schema [17] and the OMOP common data model [18] is performed. By adapting our approach to these two widely used data models a greater variety of data could be made available to process mining and conformance checking in particular.

5.2 Guideline compliance checking

In our approach we pushed the time dimension into the process structure to be able to use the conformance checking capabilities of ProM on imperative models. However, there are two viable alternative approaches: (1) Using a data-aware alignment algorithm would allow to keep the time dimension hidden in the data, thus naming the follow-up visits just *follow-up*, avoiding the initially confusing time boxing notation (e.g. I.F_00.3Q). However, we decided to use our naming approach to make all (missing) steps easily visible in the model. (2) The current version of the ProM framework also includes a declarative mining module that derives sets of constraints in form of a declarative model from log files and offers also conformance checking [13, 21]. This needs further investigation, especially in the preparation of a correct declarative constraint set based on the guideline as well as an adapted real log to be replayed.

5.3 Medical implications

In [12] the prognosis among patients with thin melanomas depending on the surveillance compliance was analyzed. Patients were considered to be compliant with the follow-up regimen if they had at least one annual follow-up examination and non-compliant if they had follow-up intervals of more than one year. They showed that compliant patients before the onset of recurrence had a significantly better prognosis than non-compliant patients.

When using our calculated fitness instead of the fixed time-intervals to evaluate the survival, the same effect can be observed in our data as seen in Figure 6. We sampled all 246 patients that stayed in follow-up for more than two years based on their fitness value into three equal-sized groups and used the Kaplan-Meier estimator for survival analysis. The survival probability of patients with

a high guideline compliance after five years is about 5% higher compared to the least compliant group. However, adding the patients that stayed for less than 2 years to the estimator, looking at all 358 patients in follow-up, showed a reversed effect. The main reason was that higher fitness is easier to achieve with a shorter stay and many with a short stay died early, e.g. after being staged in IV and the first follow-up visit.

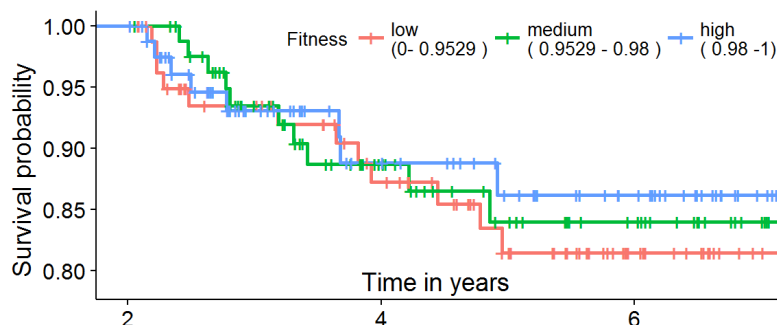


Fig. 6. Survival analysis for all 246 patients that stayed in follow-up for more than two years, sampled into three equal-sized groups depending on their fitness.

The compliance calculated and formalized using the fitness using MPE is very promising. Yet it has to be further analyzed under which circumstances it correlates to the outcome of the patients. Further we plan to analyze how the compliance affects the tumor progression of the patients, i.e. if patients with a higher compliance are less likely to progress to a higher AJCC stage.

References

1. van der Aalst, W. M.: Process Mining: discovery, conformance and enhancement of business processes. Springer Science & Business Media (2011)
2. Adriansyah, A., van Dongen, B. F., van der Aalst, W. M.: Towards robust conformance checking. In: International Conference on Business Process Management, pp. 122–133. Springer, Berlin, Heidelberg (2010)
3. Binder, M. et al.: On analyzing process compliance in skin cancer treatment: an experience report from the evidence-based medical compliance cluster (EBMC 2). In: International Conference on Advanced Information Systems Engineering, pp. 398–413. Springer, Berlin, Heidelberg (2012)
4. De Weerd, J., De Backer, M., Vanthienen, J., Baesens, B.: A critical evaluation study of model-log metrics in process discovery. In: International Conference on Business Process Management, pp. 158–169. Springer, Berlin, Heidelberg (2011)
5. Dorda, W. et al.: ArchiMed: A Medical Information and Retrieval System. *Methods Inf Med* **38**(1), 16–24 (1999)

6. Dunkl, R., Frschl, K. A., Grossmann, W., Rinderle-Ma, S.: Assessing medical treatment compliance based on formal process modeling. In: Symposium of the Austrian HCI and Usability Engineering Group, pp. 533–546. Springer, Berlin, Heidelberg (2011)
7. Francken, A. B., Bastiaannet, E., Hoekstra, H. J.: Follow-up in patients with localised primary cutaneous melanoma. *The Lancet. Oncology* **6**(8), 608–621 (2005)
8. Garbe, C.: Increasing incidence of malignant melanoma, *Hautarzt* **51**(7), p. 518 (2000). <https://doi.org/10.1007/s001050051166>
9. Garbe, C. et al.: Diagnosis and treatment of melanoma: European consensus-based interdisciplinary guideline. *European Journal of Cancer* **46**(2), 270–283 (2010)
10. Hackl, M., Ihle, P.: *Krebserkrankungen in Österreich*. 6th edn. Statistik Austria, Vienna (2018)
11. Jemal, A. et al.: Recent trends in cutaneous melanoma incidence and death rates in the United States, 1992-2006. *Journal of the American Academy of Dermatology* **65**(5), 17 (2011)
12. Kittler, H., Weitzdorfer, R., Pehamberger, H., Wolff, K., Binder, M.: Compliance with follow-up and prognosis among patients with thin melanomas. *European Journal of Cancer* **37**(12), 1504-1509 (2001)
13. Maggi, F. M.: Declarative Process Mining with the Declare Component of ProM. In: BPM (Demos), Beijing (2013)
14. Mannhardt, F., De Leoni, M., Reijers, H. A.: The Multi-perspective Process Explorer. In: BPM (Demos), Innsbruck (2015)
15. Mannhardt, F., De Leoni, M., Reijers, H. A., Van Der Aalst, W. M.: Balanced multi-perspective checking of process conformance. *Computing* **98**(4), 407–437 (2016)
16. Mans, R. S., van der Aalst, W. M., Vanwersch, R. J.: Process mining in healthcare: evaluating and exploiting operational healthcare processes. Springer (2015)
17. Murphy, S. N., Weber, G., Mendis, M., Gainer, V., Chueh, H. C., Churchill, S., Kohane, I.: Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *Journal of the American Medical Informatics Association* **17**(2), 124–130 (2010)
18. Overhage, J. M., Ryan, P. B., Reich, C. G., Hartzema, A. G., Stang, P. E.: Validation of a common data model for active safety surveillance research. *Journal of the American Medical Informatics Association* **19**(1), 54–60 (2011)
19. Rinner, C., Duftschmid, G., Wrba, T., Gall, W.: Making the complex data model of a clinical research platform accessible for teaching. *Journal of Innovation in Health Informatics* **24**(1) (2017)
20. Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: A literature review. *Journal of Biomedical Informatics* **61**, 224–236 (2016)
21. Rovani, M., Maggi, F. M., de Leoni, M., van der Aalst, W. M.: Declarative process mining in healthcare. *Expert Systems with Applications* **42**(23), 9236–9251 (2015)
22. Rozinat, A., Van der Aalst, W. M.: Conformance checking of processes based on monitoring real behavior. *Information Systems* **33**(1), 64–95 (2008)
23. Safran, C. et al.: Toward a national framework for the secondary use of health data: an American Medical Informatics Association White Paper. *Journal of the American Medical Informatics Association* **14**(1), 1–9 (2007)
24. Tax, N., Lu, X., Sidorova, N., Fahland, D., van der Aalst, W. M.: The Imprecisions of Precision Measures in Process Mining. arXiv preprint, arXiv:1705.03303 (2017)
25. Verbeek, H. M. W., Buijs, J. C., Van Dongen, B. F., Van Der Aalst, W. M.: Xes, xesame, and prom 6. In: Forum at the Conference on Advanced Information Systems Engineering (CAiSE), pp. 60–75. Springer, Berlin, Heidelberg (2010)