



# Characterization of Drug Use Patterns Using Process Mining and Temporal Abstraction Digital Phenotyping

Eric Rojas<sup>(✉)</sup> and Daniel Capurro

Internal Medicine Department, School of Medicine,  
Pontificia Universidad Católica de Chile, Santiago, Chile  
{eric.rojas,dcapurro}@uc.cl

**Abstract.** Understanding and identifying executed patterns, activities and processes for patients of different characteristics provides medical experts a deep understanding of which tasks are critical in the provided care, and may help identify ways to improve them. However, extracting these events and data for patients with complex clinical phenotypes is not a trivial task. This paper provides an approach to identifying specific patient cohorts based on complex digital phenotypes as a starting point to apply process mining tools and techniques and identify patterns or process models. Using temporal abstraction-based digital phenotyping and pattern matching, we identified a cohort of patients with sepsis from the MIMIC II database, and then apply process mining techniques to discover medication use patterns. In the case study we present, the use of temporal abstraction digital phenotyping helped us discover a relevant patient cohort, aiding in the extraction of the data required to generate drug use patterns for medications of different types such as vasopressors, vasodilators and systemic antibacterial antibiotics. For sepsis patients, combining the use of temporal abstraction digital phenotyping and process mining tools and techniques, was proven to help extract accurate cohorts of patients for health care process mining.

**Keywords:** Process mining · Digital phenotyping · Drug use patterns

## 1 Introduction

Currently, medical centers execute a large number of medical activities and procedures, which constitute an essential part of the provided care. To successfully administer the most suitable care to patients, processes must be executed in the most effective and efficient way possible.

Medications are an essential component of clinical care. The pattern at which they are administered can reflect the clinical pathways followed by patients during their episodes of care. Also, in general, following a proper medication process or pathway should also impact the healthcare system, reducing hospitalization

rates, improving quality of life, reducing health system costs, and, improving quality of care [11].

Nowadays, hospital information systems store large amounts of medical information generated during routine patient care; this includes abundant medication administration data. Using this information, it is possible to conduct several types of analysis. Data analytic tools and techniques, specifically process mining, provide the ability to discover process models, understand the interaction between resources and analyze their performance [1]. The use of process mining techniques does not only facilitate an understanding of the natural complexity of hospital processes and what these genuinely entail, but it also generates improvement opportunities in relation to care services. With the use of process mining, drug administration patterns may be studied to understand which are the drugs administered and in which patterns.

In the past, process mining in healthcare has been applied, generating complex and unreadable spaghetti process models [1]. To reduce this complexity, multiple efforts have been done. For example, applying simple filters guided by the desired objectives [23] or clustering techniques to organize patients [20]. These efforts required multiple tasks to identify the correct cohorts of patients, increasing the possibility of leaving significant patients outside of the analysis. Our study proposes the use of temporal abstraction digital phenotyping to identify more specific cohorts of patients with exact medication and condition types, to use process oriented data analytic techniques to generate detailed patterns.

The structure of the paper is as follows: Sect. 2 includes the background of the study. Section 3 describes the followed method. Section 4 describes a case study using the proposed method, including the results and discussion. Finally, conclusions and future work are included at the end of this paper.

## 2 Background

The use of clinical data is relevant in the big data era. Each time more and more data are being produced in clinical environments. In order to study the executed processes, the data must be accessed, selected and extracted faster and faster to not lose track of details included in it. The selection process of the adequate data is vital to obtain the desired results. In the past, filtering [23] and clustering [20] techniques have been good approaches to define the exact data to extract for process analysis, but can be time consuming and generate cohorts of patient's data that may not be the desired ones.

The need to generate good techniques to identify and extract data for patients satisfying very punctual conditions is required in process mining. In our approach we focus on presenting a method that combines the use of temporal abstraction digital phenotyping with the process mining tools and techniques available.

### 2.1 Temporal Abstraction Digital Phenotyping

Patient phenotyping has become a key component of complex data analytics in healthcare. The identification of predictor variables or risk factors using large

volumes of routinely collected clinical data between two patient types—for example, comparing older adults with and without dementia, understanding complex processes in emergency rooms and analyzing clinical data to identify heralding attributes—has become a prevailing study design and it requires establishing clear and robust patient phenotypes and methods to retrieve patients that match that phenotype. However, establishing robust clinical phenotypes in databases of routinely collected clinical data is not a trivial task. Current approaches include the use of billing information, diagnostic codes, developing complex algorithms, natural language processing and, ultimately, manual abstraction of clinical records [5, 13, 27]. Our research group has advanced the use of temporal abstraction and pattern matching to design digital phenotypes and query clinical databases to retrieve patients that meet the designed phenotype [4]. Briefly, there are two main approaches when attempting to identify patient cohorts based on temporal patterns of clinical data in raw medical records.

The first and probably more cumbersome one is developing a complex query using database query languages or specific query languages with temporal capacities, usually built over an existing one, such as SQL. Such an effort has been made in the past not only by clinical and health informatics researchers, but by database and data storage researchers too: temporal querying and temporal database capacities is a long dated problem. Starting in 1992, TSQL2 [26] was a significant attempt on making temporal databases a reality, until its death in 2001 after hard criticism. A decade later, SQL:2011 [25] included temporal behavior on its core definition, but with little real functionality added. This approach has emerged in some health informatics researchers as well. CHRONUS introduced TimeLineSQL (TSQL) [7], and later in CHRONUS II an extension to the SQL query language [16], greatly inspired in TSQL2 [25], was also developed as a means to directly execute temporal based queries on databases. The most significant hurdle of these methods involve developing very complex database queries that require advance programming knowledge with limited reusability.

The other main way of dealing with temporal querying difficulties is by abstracting the underlying raw data onto higher level models suitable for time expressiveness. Typically, this involves the ability to represent time concepts such as instants, intervals and bound times, along with temporal relations among them. DXtractor [15] introduced a hybrid model where simple plain SQL queries were produced to retrieve patient sets from some precompiled options, and then allowed to perform some basic Boolean and temporal operations over these sets in a chained way. IDAN [2] presents a temporal-abstraction mediation approach, implemented in a modular architecture. In a way, it decouples the effort of temporal reasoning on clinical data in several small components, each with a specific task. PROTEMPA [19] aimed to offer a system for specifying temporal and mathematical relationships between data elements to retrieve cohorts of patients that meet the given requirements. It is a framework like system, with clearly distinguishable modules for different tasks, such as data extraction, knowledge and abstractions definition, temporal trends detection, etc. A later and more novel system, Eureka [18], was constructed to address some of PROTEMPA's issues

and extend its functionality. It is commonly used to integrate with the I2B2 framework, and offers many useful tools for other integrations too. Opposed to its predecessor, it carries one general ETL (Extract, Transform and Load) system capable of extracting data from different sources, but in practice this means not only configuring some metadata XMLs (eXtensible Markup Language) to describe the underlying model, but also coding a specific data extractor to fully determine the data backend. Eureka is highly configurable and offers a rich API to integrate the researcher's database to the system, but programming knowledge and a deep understanding of how Eureka's ETL works is needed in order to achieve this. ClinicalTime [8] aims to facilitate the task of describing the database by imposing on the user no programming knowledge, only data knowledge. This means that the researcher only needs to know how his data looks like in order to define the desired mapping to the program's model. Using this knowledge the researcher is able to design a clinical phenotype and retrieve patient cohorts with detailed attributes and conditions.

## 2.2 Process Mining in Healthcare

Process mining is a research discipline that focuses on the extraction of information from data generated and stored in the databases of information systems. The data is extracted to create events logs, which can be viewed as a set of cases in which each one contains all the activities executed for a process instance [1].

Process-Aware Information Systems (PAIS) [1] are systems that should be readily able to produce event logs. Specific examples of such applications include Enterprise Resource Planning systems and Hospital Information Systems (HIS [22]). Event log data are not limited simply to the data from these tools, as many other systems can also provide useful data about process execution. Moreover, data regarding an specific complex process can come from multiple information systems or data sources.

There are three main types of process mining analysis that can be performed: process discovery, conformance checking, and enhancement. Process discovery allows process models to be extracted from an event log; conformance checking allows monitoring deviations by comparing a given model with the event log; and how enhancement allows extending or improving an existing process model using information about the actual process recorded in the event log [1].

Process mining has been successfully employed for analysis and study purposes across different industries, including the education [3], marketing [17], among others. The healthcare domain is not the exception [10, 14, 20, 22]. Normally, any activity executed in a hospital by a physician, nurse, technician or any other resource to give care to a patient is stored in a HIS (compound of databases, systems, protocols, events, etc.). Activities are recorded in event logs for support, control and further analysis. Process models are created to specify the order in which different health workers are supposed to perform their activities within a given process, or to analyze critically the process design. Moreover, process models are also used to support the development of HIS, for example, to understand how the system is expected to support the process execution [22].

### 3 Method

Figure 1 presents an overview of the methods used for the analysis. It consists of four phases that combine temporal abstraction-based digital phenotyping and process mining to discover drug patterns. The phases are the following:

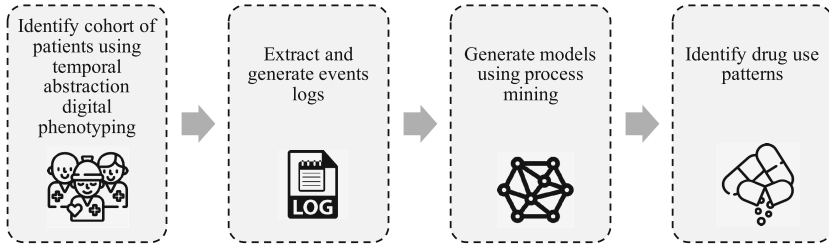


Fig. 1. Method

1. *Identify specific cohort of patients using temporal abstraction digital phenotyping from specific data sources.* Through ClinicalTime [8] a tool that includes temporal abstraction digital phenotyping, precise conditions and time intervals can be defined to extract specific patients, allowing clinicians and researchers to study more accurate groups of patients. An example of the way temporal queries are visually constructed using ClinicalTime can be seen in Fig. 2.
2. *Extract and generate events logs for the identified cohorts of patients.* Having identified an specific cohort of patients, an event log for an research purpose can be generated. An event log must include at least a case unique identifier, an executed activity, and a timestamp.
3. *Generate models using process mining tools and techniques.* With the extracted event log, several tools (such as Disco [12], or PALIA [10]) and techniques (such as heuristic miner [28] or the Palia algorithm [21]), can be used to generate process models and complementary information for analysis.
4. *Identify drug use patterns based on the models discovered in previous steps.* Using the models previously obtained, the medication data can be used to identify drug use patterns on the extracted set of identified patients.

### 4 Results

We validated the methods presented in Sect. 3 using a case study of patients with Sepsis. For each of the established steps of the method, the executed tasks are described in the next subsections.

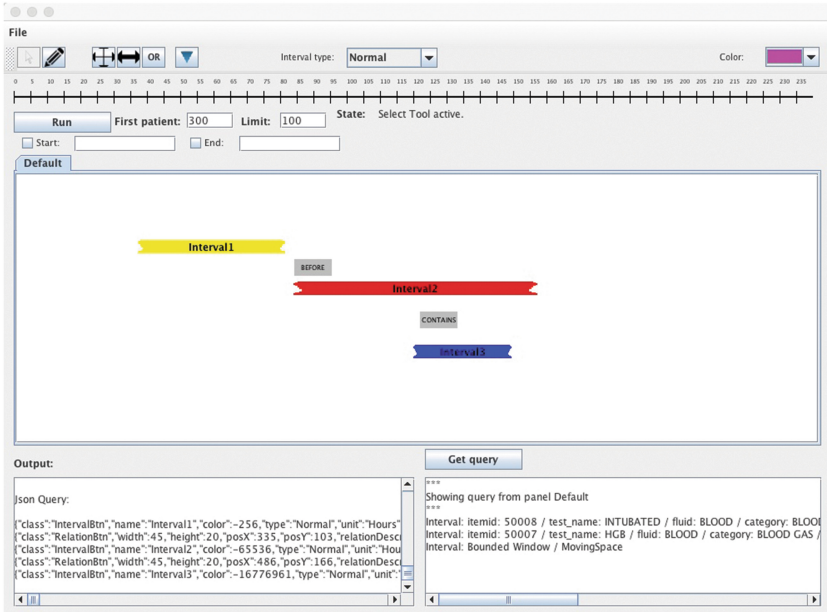


Fig. 2. Temporal query example in ClinicalTime

#### 4.1 Phase 1: Define Cohorts of Patients Using Temporal Abstraction Digital Phenotyping

For phase 1 of the method, we retrieved the patient cohort from the MIMIC II database. MIMIC II is an anonymized database containing more than 30,000 intensive care unit episodes [24]. Several tables were used to identify the cohort of patients and the data for the event log: *icustay\_events*, *ioevents*, *labevents*, *medevents*, *physicianorderentry\_events*, *procedure\_events*, *ioevents*, *microbiology\_events*, *labevents*, *medevents*, *icustay\_events*. A full description of the MIMIC II data model can be found in [6].

This database was accessed through ClinicalTime, an application to build and execute complex temporal abstraction digital phenotyping queries [8]. ClinicalTime is an application that enables researchers to define clinical phenotypes using a graphical interface. In that interface, the users define clinical temporal instants and intervals (instantaneous, bounded), as well as their temporal and mathematical relationships. Additional conditions can be defined for every interval such as: duration, number of repeated instants within an interval, etc. [8].

In addition, temporal relations can be defined between intervals and instants; ClinicalTime implements the full set of temporal relations. Researchers can define other relations between intervals such as temporal distance, an increase or decrease of a certain magnitude, a percent change in values, etc. Once an interval pattern is defined, it can be saved and combined with other interval patterns. This allows the creation of arbitrarily complex interval patterns to describe

clinical phenotypes. ClinicalTime then uses its search algorithms to abstract the temporal intervals from a clinical relational database (in our case MIMIC II), and identifies and returns a list of patients matching the pattern. This patient set becomes the patient cohort. To ensure precision, the phenotyping algorithm was validated against a manually annotated subset of MIMIC II.

For our case study we decided to extract drug use patterns for Sepsis. Sepsis is one of the main causes of admission to the intensive care unit worldwide and has significant associated morbidity and mortality. To meet the definition, a sepsis patient must meet two of the following criteria: (a) temperature  $<36^{\circ}\text{C}$  or  $>38^{\circ}\text{C}$ , (b) respiratory rate  $>20/\text{min}$  or  $\text{PaCO}_2 <32\text{ mmHg}$ , (c) heart rate  $>90/\text{min}$ , (d) white blood cell (WBC) count  $<4,000$  or  $>12,000$  or  $>10\%$  bands (immature white blood cells). In addition to these criteria, the condition must be a response to an *active infection*. An active infection was defined as a combination of clinical and laboratory results. We created the sepsis phenotype using the above criteria, and extracted from the MIMIC II a cohort of patients that meet these criteria.

## 4.2 Phase 2: Create Event Log from Identified Cohort of Patients

Based on the cohort of patients generated using Clinical Time, we proceeded to extract and generate the event log from the database MIMIC II. Based on an exploratory analysis of the data, the pharmacy provider order entry (POE) records were necessary to execute the drug use analysis.

In the database, multiple categories for medications and drugs were discovered. A drug may be classified by the chemical type of the active ingredient or by the way it is used to treat a particular condition. In our case we centered our analysis in vasodilators, vasopressors, and systemic antibacterial antibiotics. These three categories were selected to analyze. Vasodilators are medicines that dilate (widen) blood vessels, allowing blood to flow more easily through. Some act directly on the smooth muscle cells lining the blood vessels [9]. Some examples of Vasodilators are nitroglycerin and desmopressin. Vasopressors are medicines that constrict (narrow) blood vessels, increasing blood pressure. They are used in the treatment of extremely low blood pressure, especially in critically ill patients [9]. Some examples of Vasopressors are phenylephrine and dopamine. Antibiotics are drugs that can either kill an infectious bacteria or inhibit its growth. Different antibiotics work by different mechanisms and are used to treat infections caused by bacteria that are sensitive to that particular antibiotic [9]. Some examples of antibiotics are vancomycin and ampicillin.

After identifying the categories of the drugs or medications, we proceeded to extract the event logs, including all patients that have two conditions:

- (i) were identified as having Sepsis as the main cause of admission to the intensive care unit (results of phase 1), and;
- (ii) have been medicated with either vasodilators, vasopressors, or systemic antibacterial antibiotics.

Three specific cohort of patients were extracted into three event logs. For the drug use pattern of sepsis and vasodilators, 20 cases were identified with 6 start and end medication activities. For sepsis and vasopressors, 33 cases were identified with 12 start and end medication activities. And, finally, for sepsis and systemic antibacterial antibiotics, 60 cases were identified with 21.

### 4.3 Phase 3: Generate Models from Event Logs

After extracting each event log from the previous phase, process mining tools were used to generate process models. In our case, we used Disco [12] to generate the models. Following are the three resulting drug use patterns discovered.

First, the process model for patients with sepsis and vasodilators medications is presented in Fig. 3. This model includes the starting and ending medication activities for nitroglycerin, and nitroprusside sodium, which were the identified vasodilators. The arcs correspond to the sequential order for all the different 20 cases. Nitroglycerin was the most frequently used vasolidator.

Second, the process model reflecting the drug use for patients with sepsis and vasopressors medications is presented in Fig. 4. This model includes the starting and ending medication activities for norepinephrine, phenylephrine, epinephrine, vasopressin, dopamine and dobutamine, which were the identified vasopressors. The arcs correspond to the sequential order for all the different 33 cases. Norepinephrine was the vassopressor medicated the most.

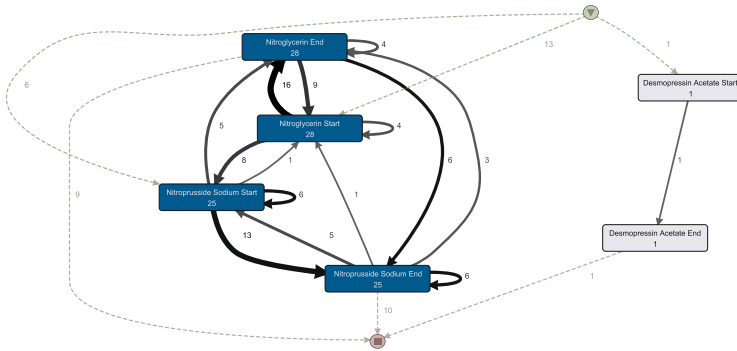
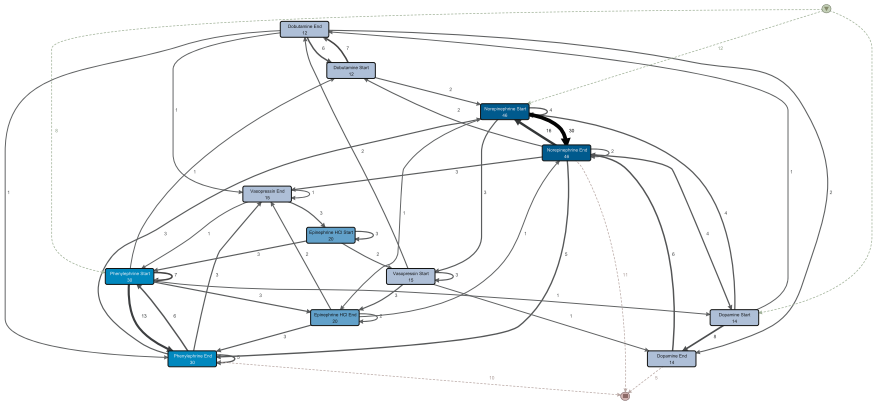


Fig. 3. Process for vasodilators

And, finally, the process model reflecting the drug use for patients with sepsis and systemic antibacterial antibiotics is presented partially in Fig. 5. This model includes only the starting medication activities for vancomycin, levofloxacin, metronidazole, piperacillin-tazobactam, ceftriaxone, aztreonam, meropenem, ampicillin, cefazolin, azithromycin, linezolid, ciprofloxacin, clindamycin, unasyn, sulfameth/trimethoprim, penicillin G, potassium, erythromycin, nafcillin, oxacillin, dicloxacillin, and, imipenem-cilastatin. These were the identified antibiotics.





**Fig. 4.** Process for vasopressors

In this case to improve the visualization of the process model, only the starting activities were selected and not the ending of the treatment. The arcs correspond to the sequential order for all the different 60 cases. Vancomycin and levofloxacin were the systemic antibacterial antibiotics most frequently prescribed.

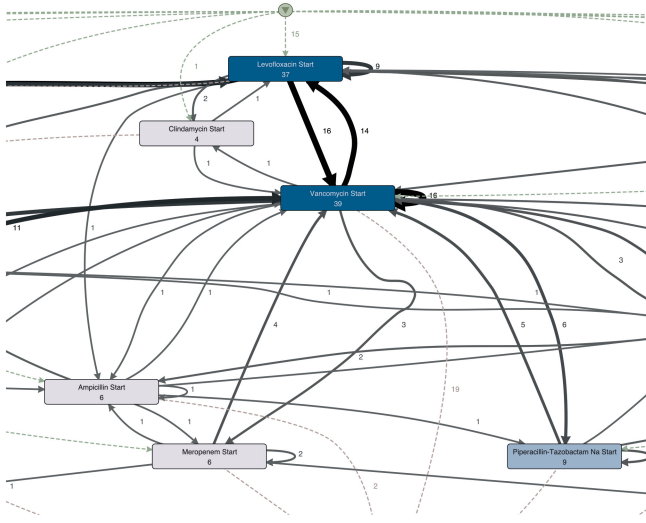
#### 4.4 Phase 4: Identify Drug Use Patterns

Finally, we identified the drug use patterns using the information and models acquired through process mining in the previous phases.

First, for the vasodilators, the main pattern is featured in Fig. 3. Although a full clinical analysis is beyond the scope of this paper, the observed pattern, its sequences and frequencies, is consistent to what clinicians use in the real-world. It is frequent, in most of the cases, to either only prescribe nitroglycerine (35%), or, prescribe it combined with nitroprusside sodium (35%), being this drug medicated in 70% of cases. Also, 65% of cases begin with nitroglycerine, being the first option as a vasodilator to be medicated.

Secondly, for the vasopressors, the main pattern is featured in Fig. 4. Twenty-four percent of all cases only prescribe Norepinephrine, while Epinephrine is never prescribed alone. The most prescribed are norepinephrine, which is present in 70% of the cases and Phenylephrine, which is present in 42% of them. Additional to this, 27% of the cases include two different vasopressors, and only 12% include 4 or more different ones. The trend in these cases is to start either with Norepinephrine in 42% of the cases, or start with Phenylephrine in 30%.

And finally, for the systemic antibacterial antibiotics, part of the main pattern is shown in Fig. 5, where some characteristics were identified. There is no predominant antibiotic that is prescribed alone, the top antibiotics prescribed alone are levofloxacin (only 8.3%) and cefazolin (only 5%). More than 50% of the cases prescribe 3 or more antibiotics (53.3%), while 20% prescribe only two. Fifty-three percent of the cases tend to start the antibiotic medication pattern



**Fig. 5.** Process for antibiotics

with either vancomycin (28.3%) or levofloxacin (25%), being these two the initials and the most prescribed antibiotics. While no cases start directly with aztreonam, linezolid, nafcillin, oxacillin or dicloxacillin. When clinicians are presented with these patterns, they are concordant with current clinical practices, highlighting the potential of the presented methods for clinical data analytics.

## 5 Discussion

The discussion will be addressed both from a process mining and a clinical perspective. From the process mining point of view our method and case study contributed with an easier and more accurate method to identify cohorts of patients when multiple and temporally-related clinical conditions must be met. This facilitated the extraction and generation of event logs, because only the necessary data was included for the analysis, eliminating or reducing the filtering/clustering phases utilized by most process-mining methodologies. In addition, this simplified the complexity involved in generating accurate queries to extract data from electronic health records. Finally, this approach generated better, more understandable and readable models, that are based on cohorts of patients with very specific conditions, making them easier to analyze.

From the clinical point of view, analyzing drug use patterns may help in measuring conformance of clinical practice with guideline recommendations, identify changes in prescription pattern over time, for example, when new resources or drugs are incorporated to care and whether those changes are clinically explained or not. This method also generates ways to verify, monitor and control drug prescription on specific groups of patients, opening avenues to improving the provided care and the quality of the outcome.

## 6 Conclusions and Future Work

In this case study, we applied a method based on the combination of process mining and temporal abstraction-based digital phenotyping to help discover drug use patterns among patients with sepsis. Each step of the proposed method has been executed: identifying cohorts of patients using temporal abstraction digital phenotyping, creating the event logs, generating models, and finally identifying drug use patterns. Although beyond the scope of this demonstration, the ability to quickly discover drug utilization patterns should be a useful tool to study healthcare resource utilization, its patterns and how they might change over time. Future work will include expanding the case study to include additional data sources and phenotypes (such as additional drug classes and clinical outcomes), the inclusion of additional process mining techniques to improve pattern discovery, and, finally, formal clinical validation of the identified patterns.

## References

1. 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>
2. Boaz, D., Shahar, Y.: Idan: a distributed temporal-abstraction mediator for medical databases. In: Dojat, M., Keravnou, E.T., Barahona, P. (eds.) *AIME 2003*. LNCS (LNAI), vol. 2780, pp. 21–30. Springer, Heidelberg (2003). [https://doi.org/10.1007/978-3-540-39907-0\\_3](https://doi.org/10.1007/978-3-540-39907-0_3)
3. Bogarín, A., Romero, C., Cerezo, R., Sánchez-Santillán, M.: Clustering for improving educational process mining. In: *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*, pp. 11–15. ACM (2014)
4. Capurro, D., Barbe, M., Daza, C., Santa María, J., Trincado, J., Gomez, I.: *ClinicalTime: identification of patients with acute kidney injury using temporal abstractions and temporal pattern matching*. *AMIA Summits Transl. Sci. Proc.* **2015**, 46 (2015)
5. Che, Z., Kale, D., Li, W., Bahadori, M.T., Liu, Y.: Deep computational phenotyping. In: *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 507–516. ACM (2015)
6. Clifford, G.D., Scott, D.J., Villarroel, M., et al.: User guide and documentation for the MIMIC II database. *MIMIC-II database version 2(95)* (2009)
7. Das, A.K., Tu, S.W., Purcell, G.P., Musen, M.A.: An extended SQL for temporal data management in clinical decision-support systems. In: *Proceedings of the Annual Symposium on Computer Application in Medical Care*, p. 128. American Medical Informatics Association (1992)
8. Daza, C., Santa Maria, J., Gomez, I., Barbe, M., Trincado, J., Capurro, D.: Phenotyping intensive care unit patients using temporal abstractions and temporal pattern matching. In: *Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, pp. 508–509. ACM (2016)
9. *Drugs.com: Date source of drug information online* (2018). <https://www.drugs.com>. 1 May 2018
10. 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)

11. Gilberg, K., Laouri, M., Wade, S., Isonaka, S.: Analysis of medication use patterns: apparent overuse of antibiotics and underuse of prescription drugs for asthma, depression, and CHF. *J. Manag. Care Pharm.* **9**(3), 232–237 (2003)
12. Günther, C., Rozinat, A.: Disco: discover your processes. *BPM* **940**, 40–44 (2012)
13. Liao, K.P., et al.: Development of phenotype algorithms using electronic medical records and incorporating natural language processing. *BMJ* **350**, h1885 (2015)
14. Mans, R.S., van der Aalst, W.M.P., Vanwersch, R.J.B.: *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*. SBPM. Springer, Cham (2015). <https://doi.org/10.1007/978-3-319-16071-9>
15. Nigrin, D.J., Kohane, I.S.: Temporal expressiveness in querying a time-stamp-based clinical database. *JAMIA* **7**(2), 152–163 (2000)
16. O'Connor, M.J., Tu, S.W., Musen, M.A.: The Chronus II temporal database mediator. In: *Proceedings of the AMIA Symposium*, p. 567 (2002)
17. Osses, A.S., et al.: Business process analysis in advertising: an extension to a methodology based on process mining projects. In: *2016 35th International Conference of the Chilean Computer Science Society (SCCC)*, pp. 1–12. IEEE (2016)
18. Post, A.R., et al.: Temporal abstraction-based clinical phenotyping with Eureka! In: *AMIA Annual Symposium Proceedings*, vol. 2013, p. 1160. American Medical Informatics Association (2013)
19. Post, A., Harrison Jr., J.: Protempa: a method for specifying and identifying temporal sequences in retrospective data for patient selection. *J. Am. Med. Inform. Assoc.* **14**(5), 674–683 (2007)
20. Rebuge, Á., Ferreira, D.R.: Business process analysis in healthcare environments: a methodology based on process mining. *Inf. Syst.* **37**(2), 99–116 (2012)
21. Rojas, E., et al.: Palia-er: bringing question-driven process mining closer to the emergency room. In: *Proceedings of the Business Process Management Conference 2018 (BPM 2018) Demo Track*, Barcelona, Spain (2018)
22. Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: a literature review. *J. Biomed. Inform.* **61**, 224–236 (2016)
23. Rojas, E., Sepúlveda, M., Munoz-Gama, J., Capurro, D., Traver, V., Fernandez-Llatas, C.: Question-driven methodology for analyzing emergency room processes using process mining. *Appl. Sci.* **7**(3), 302 (2017)
24. Saeed, M., et al.: Multiparameter intelligent monitoring in intensive care II (MIMIC-II): a public-access intensive care unit database. *Crit. Care Med.* **39**(5), 952 (2011)
25. Snodgrass, R.T.: *The TSQL2 Temporal Query Language*, vol. 330. Springer, New York (2012)
26. Snodgrass, R.: *The TSQL2 Temporal Query Language*. Springer, London (1995). <https://doi.org/10.1007/978-1-4615-2289-8>
27. Wei, W.Q., Teixeira, P.L., Mo, H., Cronin, R.M., Warner, J.L., Denny, J.C.: Combining billing codes, clinical notes, and medications from electronic health records provides superior phenotyping performance. *JAMIA* **23**(e1), e20–e27 (2015)
28. Weijters, A.J.M.M., Ribeiro, J.T.S.: Flexible heuristics miner (FHM). In: *2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)* (2011)