Adopting Standard Clinical Descriptors for Process Mining Case Studies in Healthcare

Emmanuel Helm[1,3][0000−0002−1323−3511], Anna M. Lin[1][0000−0003−4840−8788], David Baumgartner[1][0000−0002−0189−4718], Alvin C. Lin[2], and Josef Küng[3]

[1] University of Applied Sciences Upper Austria, Research Department Advanced Information Systems and Technology, 4232 Hagenberg, Austria
{firstname.lastname}@fh-hagenberg.at
[2] University of Toronto, Faculty of Medicine, Toronto, Canada
alvin.lin@mail.utoronto.ca
[3] Johannes Kepler University, Institute for Applied Knowledge Processing, 4040 Linz, Austria josef.kueng@jku.at

Abstract. Process mining can provide greater insight into medical treatment processes and organizational processes in healthcare. A review of the case studies in the literature has identified several different common aspects for comparison, which include methodologies, algorithms or techniques, medical fields and healthcare specialty. However, from a medical perspective, the clinical terms are not reported in a uniform way and do not follow a standard clinical coding scheme. Further, the characteristics of the event log data are not always described. In this paper, we identified 38 clinically-relevant case studies of process mining in healthcare published from 2016 to 2018 that described the tools, algorithms and techniques utilized, and details on the event log data. We then assigned the clinical aspects of patient encounter environment, clinical specialty and medical diagnoses using the standard clinical coding schemes SNOMED CT and ICD-10. The potential outcomes of adopting a standard approach for describing event log data and classifying medical terminology using standard clinical coding schemes are discussed.

Keywords: Process Mining · Healthcare · Terminology · ICD · SNOMED.

1 Introduction

Process mining is a discipline that allows for greater understanding into real-life processes of recorded systems behaviour. Through process mining techniques, numerous case studies and successful companies have demonstrated valuable insights into quality improvement, compliance, and optimization of processes.

In recent years, several review papers provide an overview on the state of process mining in healthcare. Rojas et al in 2016 identified classifiers of eleven common aspects across 74 case studies in healthcare [35]. These aspects include methodologies, techniques or algorithms, medical fields and healthcare specialty. In 2018, Erdogan and Tarhan conducted a systematic mapping of 172 case studies with mostly the same metrics and aspects [14]. These papers are very specific

as to *how* these case studies were conducted, which enhances comparison between different process mining techniques in different settings. However, from a medical perspective, the terms and categories listed under *medical fields* and *healthcare specialty* are not structured in a uniform way, and do not follow a standardized clinical coding scheme. Further, basic characteristics of the event log data (timeframe, number of cases or patients, healthcare facility/organization) are not always clearly reported.

The number of case studies on process mining in healthcare continues to increase steadily. As such, a standard approach of reporting event log data, clinical specialties and medical diagnoses would provide greater clarity and enhance comparability between treatments of specific diseases across different heathcare settings.

In this paper, further to the studies examined by Rojas et al., we conducted a forward search of processing mining case studies in healthcare for the three-year period from January 2016 to December 2018. We identified case studies that described basic characteristics of the event log data, and where information on the patient encounter environment, clinical specialty and medical diagnoses could be assigned under a standard clinical coding scheme. Section 2 describes how the forward search was conducted and which criteria we applied to filter the results. In addition, the methods describe standard clinical coding systems and terminologies that were used. In Section 3, the results of our analysis are presented. Section 4 discusses the benefits and gives an outlook on the potential clinical insights gained by reporting and classifying clinical terms, clinical specialties and medical diagnoses using a standard clinical coding scheme.

2   Methods

Our paper focused on answering three questions: (1) Which clinically-relevant case studies of process mining in healthcare will be selected for this study? (2) What were the technical aspects identified? (3) How can we improve the clarity and comparability of the clinical terms and aspects described?

2.1   Selection of clinically-relevant case studies

Our starting point was the review paper by Rojas et al. [35] that identified 74 case studies where process mining tools, techniques or algorithms were applied in the healthcare domain. We then performed a forward search using Google Scholar, in reference to the 74 identified articles and the review paper itself. The inclusion criteria (IC) were applied in the Google Scholar search and the exclusion criteria (EC) were applied manually afterwards (see Figure 1).

IC1: All articles that reference either the review paper by Rojas et al. [35] or any of the 74 articles identified in their review were included.

IC2: All articles published between 01.01.2016 and 31.12.2018 were included.

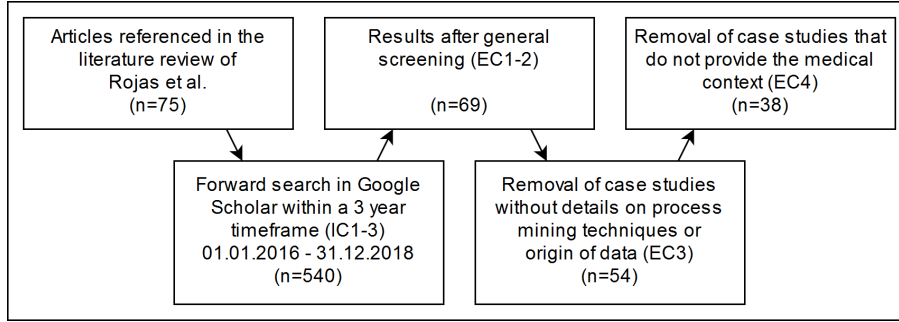**IC3:** All articles published in English were included.

Fig. 1: Flowchart on the case study selection strategy.

EC1: Articles that do not include evidence of a clinically-relevant case study of process mining in healthcare were excluded.

EC2: Articles that present a case study based on data that was already used for an earlier case study were excluded.

EC3: Articles that do not describe the characteritics of the event log data (e.g. timeframe, number of cases or patients, healthcare facility) or do not describe which process mining technique or algorithm was applied were excluded.

EC4: Articles that did not describe any clinical context (i.e. clinical specialty or medical diagnosis) were excluded.

2.2 Technical aspects

A detailed account of the tools, techniques or algorithms used in process mining case studies in healthcare have been previously described [35]. Also, other technical descriptors such as the data type and geographical analysis have been used to describe the event log data [34]. For the technical scope of our paper, our focus was on (1) the *tools* used in the case studies, (2) the *techniques or algorithms* used, and (3) the *process mining perspectives*.

2.3 Clinical aspects and standard coding schemes

Medical language is full of homonyms, synonyms, eponyms, acronyms and abbreviations; and each healthcare specialty comes with their own sub-terminology [7]. To improve the clarity and comparability of the clinical aspects described in our selected papers, we adopted the use of standard clinical coding schemes of SNOMED CT and ICD-10. Namely, the clinical terms were matched to their best corresponding standard clinical descriptor, with respect to three clinical categories: (1) the type of patient encounter environment (2) clinical speciality and (3) medical diagnosis (i.e. disease or health problem).

SNOMED CT The Systematized Nomenclature of Medicine – Clinical Terms is an internationally recognized standard that classifies clinically-relevant terminology and concepts, along with their synonyms and relationships, into numeric coded values. Available in multiple languages and maintained by SNOMED International, there are currently over 340,000 numerically coded concepts that can be combined grammatically to create an expression. We used SNOMED CT international browser[4] in version *v20190131* for clinical descriptors on the *Patient encounter environment* and *Clinical specialty*.

ICD-10 For classification of clinical diagnoses and health problems, the commonly accepted system is the International Classification of Diseases or ICD, which is maintained by the World Health Organization (WHO). The most current version is ICD-10 and it utilizes an alphanumeric coding scheme with more than 14.000 single clinical codes of medical terms organized hierarchically into 22 chapters. We used the WHO ICD-10 browser in the *2016* version[5] for clinical descriptors on *medical diagnoses*.

3   Results

3.1   Selection of clinically-relevant case studies

Our forward search yielded initially a total of 540 papers, and after our inclusion and exclusion criteria were applied, 38 articles were selected (cf. Figure 1). For all 38 papers, basic characteristics of the event log data were retrieved (e.g. origin of data, number of cases or patients, healthcare facility, timeframe of the study). The results of the technical and clinical aspects are described below.

3.2   Technical aspects

Tools ProM[6] is the most used tool in our 38 papers [1, 3, 4, 9, 16, 17, 23, 25–28, 33, 37, 39, 40, 43](n=16). Additionally, Disco[7] is also frequently used [2, 13, 16, 24, 26, 29, 33, 34, 36, 38, 44](n=11). Tools like PALIA [10, 15], pMineR [18] and others [4, 5, 8, 10, 11, 15, 31, 32, 42](n=9) are less frequently used. Four papers presented new self-developed solutions [20–22, 31].

Techniques or algorithms Fuzzy miner (which is also implemented in Disco in a next-generation version) is the most utilized algorithm [2, 13, 16, 24, 26, 29, 33, 34, 36, 38, 44](n=11). Many papers also presented self-developed approaches [1, 9, 11, 12, 19–21, 27, 43, 44](n=10), with most of the approaches implemented within the ProM framework. Four studies used the Trace Clustering technique [8, 17, 24, 31]. While the Heuristic Miner algorithm was frequented as per previous reviews, [14, 35], it was only used in one of our 38 selected papers [25].

---

[4] `https://browser.ihtsdotools.org/`
[5] `https://icd.who.int/browse10/2016/en`
[6] `https://www.promtools.org`
[7] `https://fluxicon.com/disco/`

Process mining perspectives The analysis showed that 30 of the total 38 case studies mainly aimed for the *Control Flow* perspective in their dataset [1, 3–5, 8, 9, 11, 12, 15–17, 19–21, 23, 24, 27–29, 31, 32, 34, 36–40, 42–44]. Five papers analyzed the *Conformance* perspective [18, 22, 25, 26, 33], two focused on *Organizational* [2, 10] and one on *Performance* [13].

3.3   Clinical aspects using standard clinical descriptors

Encounter Environment From the patient's perspective, we considered five clinical settings or encounter environments: (1) Inpatient, (2) Outpatient, (3) Accident and Emergency department or AED, (4) General practitioner or GP practice site (5), and Pharmacy. All five encounter environments could be coded by SNOMED CT. For each paper, at least one of these five encounter environments was retrieved. Most of the papers examined events within the Inpatient environment, followed by AED environment (cf. Table 1).

SNOMED CT	Environment	Papers
440654001	Inpatient	[1, 4, 5, 8, 9, 11, 13, 17, 19–25, 27–29, 31–34, 36–40, 42–44]
440655000	Outpatient	[3, 5, 16, 18, 26]
225728007	AED	[2, 3, 12, 15, 27, 28, 32, 36, 40, 42–44]
394761003	GP practice site	[10]
264372000	Pharmacy	[34]

Table 1: Papers with their corresponding SNOMED CT encounter environment.

Clinical Specialty SNOMED CT offers the code of 394658006 for *Clinical specialty*, which further contains 18 high-level specialties. Table 2 shows 11 of the 18 high-level clinical specialties were identified in our selected papers. The most identified clinical specialty was *Medical specialty*, followed by *Surgical specialty* and *Emergency medicine*. Of note, some of 18 high-level specialties in SNOMED CT are further divided into sub-specialties of greater clinical specificity. For example, *Medical specialty* has 44 sub-specialties that include e.g. *Dermatology*, *Neurology* and *Cardiology*. In this paper, we identified and assigned sub-specialities to their corresponding high-level *Clinical specialty*. Also, for example, if several different medical sub-specialities were described in one paper, we counted these sub-specialities together as *Medical specialty*.

Medical Diagnosis For each paper, we focused on identifying the medical diagnosis (i.e. disease or health problem) or description of a medical diagnosis. We then assigned these terms to their corresponding highest chapter or block category in ICD-10. Table 3 shows a total of 15 out of the 22 ICD-10 chapter categories for disease and health related problems were covered amongst the papers. The category with the most papers listed was *Diseases of the circulatory*

SNOMED CT	Clinical Specialty	Papers
394592004	Clinical oncology	[5, 11, 39]
394581000	Community medicine	[10]
722163006	Dentistry	[16, 26]
722164000	Dietetics and nutrition	[10]
773568002	Emergency medicine	[2–4, 12, 15, 27, 28, 32, 36, 40, 42, 43]
394814009	General practice	[10, 16, 26]
408446006	Gynecological oncology	[21]
394733009	Medical specialty	[1, 4, 8, 9, 17–20, 22, 25, 26, 29, 31, 33, 34, 38, 40, 42]
722165004	Nursing	[2, 10, 26]
394585009	Obstetrics and gynecology	[9, 21]
394732004	Surgical specialty	[4, 11, 14, 18, 21, 23, 24, 26, 37, 39, 43]

Table 2: Papers with their corresponding SNOMED CT clinical specialty.

system followed by *Neoplasms*. Two papers [9, 25] were not included in Table 3, since several hundred diseases and health problems were cited and classified using ICD-9. Of the remaining 36 case studies, ICD-10 was already used in 8 papers to code the diagnosis [4, 5, 8, 22, 25, 31, 39, 40].

4   Discussion

Whether for process discovery, conformance checking, or enhancement, process mining case studies are influenced by the quality of the labeled data. The benefits of high-quality, labeled data include improved accuracy, efficiency and predictability of processes, not only for the study itself but also for comparability across studies. Further, high-quality, labeled data can make other kinds of future analyses and even machine learning techniques (e.g. supervised learning, trend estimation, clustering, ...) easier and more efficient to achieve. In process mining case studies in healthcare, labeled data often encompasses clinical aspects and terms. As such, our aim was to examine clinically-relevant case studies since Rojas et al [34] and determine how to improve upon the clarity and comparability of clinical aspects and terms described.

4.1   Reporting basic characteristics of the event log data

For our analysis, we selected papers that described basic characteristics of the event log data. These characteristics included the origin or source of the data, the healthcare facility, the number of cases or patients, and the timeframe of the study. For example, in Rinner et al. [33], event logs were extracted for a total of 1023 patients starting melanoma surveillance between January 2010 to June 2017, from a local melanoma registry in a medical university and Hospital Information System (HIS) in Austria. In papers where these characteristics were not

ICD-10	Diagnosis	Papers
A00 - B99	Certain Infectious and parasitic diseases	[4, 27, 28, 32, 34, 40]
C00 - D48	Neoplasms	[4, 5, 11, 21, 29, 33, 37, 39]
E00 - E90	Endocrine, nutritional and metabolic diseases	[1, 4, 10, 18, 19, 26]
F00 - F99	Mental and behavioural disorders	[4, 22]
G00 - G99	Diseases of the nervous system	[4]
H60 - H95	Diseases of the ear and mastoid process	[4]
I00 - I99	Diseases of the circulatory system	[1, 4, 8, 13, 15, 17, 19, 20, 29, 31, 42]
J00 - J99	Diseases of the respiratory system	[2, 4]
K00 - K93	Diseases of the digestive system	[2, 4, 16, 36]
M00 - M99	Diseases of the musculoskeletal system and connective tissue	[2, 4, 22]
N00 - N99	Diseases of the genitourinary system	[4]
O00 - O99	Pregnancy, childbirth and the puerperium	[4]
R00 - R99	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	[2, 4, 12]
S00 - T98	Injury, poisoning and certain other consequences of external causes	[2–4, 12, 38, 43, 44]
Z00 - Z99	Factors influencing health status and contact with health services	[4, 11, 18, 23, 24]

Table 3: Papers with their corresponding ICD-10 medical diagnosis.

clearly reported, the retrieval process was time-consuming. Several papers provided additional details (e.g. patient age, data from private insurance or public health records). Presumably for reasons of privacy and anonymity, specifics on the healthcare facility (e.g. hospital name) were not always provided, however, the country of origin was always reported. While variations exist in the style of reporting, we recommend case studies include these aforementioned basic characteristics when reporting the event log data.

4.2 Adopting the use of standard clinical descriptors

Encounter Environment A patient can have vastly different experiences within the healthcare system depending on the clinical setting or encounter environment. For example, a patient with heart failure who presents to the AED may require admission as a hospital inpatient, follow-up at their GP practice site or outpatient clinic, and prescription drugs at a pharmacy. As such, in our analysis of the selected papers, we focused on five patient encounter environments: Inpatient, Outpatient, AED, GP practice site, and Pharmacy. All five encounter types can be coded by SNOMED CT. While further details can be provided (e.g. Outpatient Clinic for Thyroid Disease [18]), we recommend case studies report at least the patient encounter environment using standard clinical codes e.g. SNOMED CT.

Clinical Specialty Different clinical specialties are often involved in the care of a patient. For example, for a patient diagnosed with cancer, a multidisciplinary care plan can encompass input from a medical specialty, a surgical specialty and clinical oncology. As each specialty offers their own unique set of knowledge and expertise, it is important to identify which clinical specialty is involved.

For each of our selected papers, we identified at least one of the 18 high-level clinical specialties coded by SNOMED CT. For greater specificity, SNOMED CT offers further standard clinical codes for sub-specialities. In fact, Baek et al. list multiple sub-specialities along with their corresponding SNOMED CT codes in their study [4]. Also, instead of *Clinical specialty*, another category of clinical descriptors such as the type of medical practitioner or occupation could have been considered (e.g. mapping to surgeon instead of surgical specialty).

In any event, the task of identifying and assigning such standard clinical codes is time consuming, and beyond the scope of this paper. For future case studies, we recommend reporting the clinical specialty (or similar clinical descriptor such as medical practitioner) by adopting standard clinical codes e.g. SNOMED CT.

Medical Diagnosis There are literally thousands of medical diagnoses, and each diagnosis comes with its own treatment and management plan. ICD-10 is a standard coding scheme in healthcare that provides specific clinical descriptors and codes for diseases and health conditions. In our analysis, we were able to identify at least one medical diagnosis or description of a medical diagnosis in each paper, which we could map to the corresponding ICD-10 code. Further, over 25% (10 out of 38) of our selected papers utilized either ICD-9 or ICD-10 codes in their study. For broader comparison across studies, we assigned the selected papers to one or more of the 22 ICD-10 chapters or block categories. In Table 3 we only listed the ICD-10 chapters that were covered in the case studies.

It is important to distinguish the difference between a medical diagnosis (i.e. the process of identifying the disease or medical condition that explains a patient's signs and symptoms) versus a patient's signs (e.g. rash) or symptoms (e.g. cough). While the majority of ICD-10 chapters describe a group of medical diagnoses, some cover other clinical descriptors, such as signs and symptoms (R00-R99), external causes of morbidity and mortality (V01-V98), and codes for special purposes (U00-99). ICD-10 also allows for the coding of location, severity, cause, manifestation and type of health problem [41].

Taken together, we recommend adopting use of a standard coding scheme e.g. ICD-10 for clinical terms and aspects relating to medical diagnosis in process mining case studies in healthcare. Recently developed, ICD-11 is not adopted yet but provides backward compatibility, i.e. ICD-10 coded case studies will be comparable to newer ICD-11 coded ones, once the new coding scheme will be taken on by the information system vendors.

4.3   Conclusions and Future Perspectives

In summary, we propose adopting a standard for describing event log data and reporting medical terminology using standard clinical descriptors and coding

schemes. In scientific research, the idea of having a set of guidelines, criteria, or standards for peer-reviewed publications is not novel. In fact, journals such as *Nature* are taking initiatives by creating mandatory reporting summary templates[8], in order to improve comparability, transparency, and reproducibility of the work they publish [30]. Other journals and disciplines, including biomedical informatics, are following suit [6]. Thus, as data sets become more transparent and available, consistency in reporting characteristics of the event log data (e.g. origin of data, number of patients or cases, healthcare facility, timeframe of the study) will aid in improving comparability and reproducibility across future studies.

Further to the work by Rojas et al. [36], we identified and described the clinical terms and aspects in our selected papers with respect to three categories: the patient encounter environment, clinical specialty, and medical diagnosis. We then correlated the clinical terms and aspects to their respective standard clinical descriptors and codes found in SNOMED CT and ICD-10. For each of the five types of patient encounter environments, a more detailed description can be achieved through SNOMED CT by using compositional grammar. Similarly, for *Clinical specialty* in SNOMED CT, reporting of sub-specialties under e.g. *Medical speciality* will provide increased specificity for clarity and comparison.

As aforementioned, several case studies have already adopted the use of a standard clinical coding scheme to describe medical diagnoses. However, our consideration of SNOMED CT and ICD-10 serves only as a starting point. In fact, SNOMED CT also provides standard codes for medical diagnoses, which can provide further specificity and clarity. For example, instead of ICD-10, the one of Systematized Nomenclature for Dentistry or SNODENT CT (which is incorporated into SNOMED CT) could have been used to code for the clinical descriptors of missing and filled teeth in one of our selected papers [16].

Finally, when adopting the use of standard clinical descriptors, we recognize other fundamental clinical categories to consider are medical investigations and procedures. As such, the use of standard clinical descriptors is becoming increasingly relevant, not only for clarity and comparability, but efficiency in outcome measurements such as length of stay (LOS) and financial cost. For example, in their paper, Baek et al. utilized process mining techniques and statistical methods to identify the factors associated with LOS in a South Korean hospital [4]. This study is just one use case for a more detailed description of the medical context where process mining case studies could allow for future meta-studies, e.g. benchmarking LOS in different hospitals or countries, based on diagnoses while also considering other important factors like the patient encounter environment.

---

[8] `https://www.nature.com/documents/nr-reporting-summary-flat.pdf`

Bibliography

[1] Alharbi, A., Bulpitt, A., Johnson, O.: Improving Pattern Detection in Healthcare Process Mining Using an Interval-Based Event Selection Method. In: International Conference on Business Process Management. pp. 88–105. Springer (2017)

[2] Alvarez, C., Rojas, E., Arias, M., Munoz-Gama, J., et al.: Discovering role interaction models in the Emergency Room using Process Mining. Journal of biomedical informatics **78**, 60–77 (2018)

[3] Andrews, R., Wynn, M.T., Vallmuur, K., ter Hofstede, A.H.M., Bosley, E., Elcock, M., Rashford, S.: Pre-hospital retrieval and transport of road trauma patients in Queensland: A process mining analysis (2018)

[4] Baek, H., Cho, M., Kim, S., Hwang, H., Song, M., Yoo, S.: Analysis of length of hospital stay using electronic health records: A statistical and data mining approach. PloS one **13**(4) (2018)

[5] Baker, K., Dunwoodie, E., et al.: Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy. International journal of medical informatics **103**, 32–41 (2017)

[6] Bakken, S.: The journey to transparency, reproducibility, and replicability. Journal of the American Med. Informatics Association **26**, 185–187 (2019)

[7] Benson, T., Grieve, G.: Principles of health interoperability: SNOMED CT, HL7 and FHIR. Springer (2016)

[8] Chen, J., Sun, L., Guo, C., Wei, W., Xie, Y.: A data-driven framework of typical treatment process extraction and evaluation. Journal of biomedical informatics **83**, 178–195 (2018)

[9] Chen, Y., Kho, A.N., Liebovitz, D., Ivory, C., Osmundson, S., Bian, J., Malin, B.A.: Learning bundled care opportunities from electronic medical records. Journal of biomedical informatics **77**, 1–10 (2018)

[10] Conca, T., Saint-Pierre, C., Herskovic, V., Sepúlveda, M., Capurro, D., Prieto, F., Fernandez-Llatas, C.: Multidisciplinary Collaboration in the Treatment of Patients With Type 2 Diabetes in Primary Care: Analysis Using Process Mining. Journal of medical Internet research **20**(4) (2018)

[11] Dagliati, A., Sacchi, L., Zambelli, A., Tibollo, V., Pavesi, L., Holmes, J.H., Bellazzi, R.: Temporal electronic phenotyping by mining careflows of breast cancer patients. Journal of biomedical informatics **66**, 136–147 (2017)

[12] Duma, D., Aringhieri, R.: An ad hoc process mining approach to discover patient paths of an Emergency Department. Flexible Services and Manufacturing Journal pp. 1–29 (2018)

[13] Erdogan, T.G., Tarhan, A.: A Goal-Driven Evaluation Method Based On Process Mining for Healthcare Processes. Applied Sciences **8**(6), 894 (2018)

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

[15] Fernandez-llatas, C., Ibanez-sanchez, G., et al.: Analyzing Medical Emergency Processes with Process Mining : The Stroke Case. In: International Conference on Business Process Management. pp. 214–225. Springer (2018)

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

[17] Funkner, A.A., Yakovlev, A.N., Kovalchuk, S.V.: Data-driven modeling of clinical pathways using electronic health records. Procedia Computer Science **121**, 835–842 (2017)

[18] Gatta, R., Vallati, M., Lenkowicz, J., Casa, C., Cellini, F., Damiani, A., Valentini, V.: A Framework for Event Log Generation and Knowledge Representation for Process Mining in Healthcare. In: International Conference on Tools with Artificial Intelligence (ICTAI). pp. 647–654. IEEE (2018)

[19] Huang, Z., Dong, W., Ji, L., He, C., Duan, H.: Incorporating comorbidities into latent treatment pattern mining for clinical pathways. Journal of Biomedical Informatics **59**, 227–239 (2016)

[20] Huang, Z., Ge, Z., Dong, W., He, K., Duan, H.: Probabilistic modeling personalized treatment pathways using electronic health records. Journal of Biomedical Informatics **86**, 33–48 (2018)

[21] Jimenez-Ramirez, A., Barba, I., Reichert, M., Weber, B., Del Valle, C.: Clinical Processes-The Killer Application for Constraint-Based Process Interactions? In: International Conference on Advanced Information Systems Engineering. pp. 374–390. Springer (2018)

[22] Johnson, O., Dhafari, T.B., Kurniati, A., Fox, F., Rojas, E.: The clearpath method for care pathway process mining and simulation. In: International Conference on Business Process Management. pp. 239–250. Springer (2018)

[23] Kirchner, K., Marković, P.: Unveiling Hidden Patterns in Flexible Medical Treatment Processes A Process Mining Case Study. In: Int. Conference on Decision Support System Technology. pp. 169–180. Springer (2018)

[24] Kirchner, K., Marković, P., Delias, P.: Automatic creation of clinical pathways - a case study. Data Science and Business Intelligence **179**, 188 (2016)

[25] Kurniati, A.P., Rojas, E., Hogg, D., Hall, G., Johnson, O.: The assessment of data quality issues for process mining in healthcare using Medical Information Mart for Intensive Care III, a freely available e-health record database. Health Informatics Journal (2018)

[26] Lismont, J., Janssens, A.S., Odnoletkova, I., et al.: A guide for the application of analytics on healthcare processes: A dynamic view on patient pathways. Computers in Biology and Medicine **77**, 125–134 (2016)

[27] Mannhardt, F., Blinde, D.: Analyzing the trajectories of patients with sepsis using process mining. CEUR Workshop Proceedings **1859**, 72–80 (2017)

[28] Mannhardt, F., Toussaint, P.J.: Revealing Work Practices in Hospitals Using Process Mining. Building Continents of Knowledge in Oceans of Data: The Future of Co-Created eHealth (2018)

[29] Metsker, O., Yakovlev, A., Bolgova, E., Vasin, A., Koval-chuk, S.: Identification of Pathophysiological Subclinical Variances During Complex Treatment Process of Cardiovascular Patients. Procedia computer science **138**, 161–168 (2018)

[30] Munafò, M.R., Nosek, B.A., Bishop, D.V.M., et al.: A manifesto for reproducible science. Nature Human Behaviour **1**(1), 21 (2017)

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

[32] Neira, R.A.Q., de Vries, G.J., Caffarel, J., Stretton, E.: Extraction of Data from a Hospital Information System to Perform Process Mining. In: MedInfo. pp. 554–558 (2017)

[33] Rinner, C., Helm, E., Dunkl, R., Kittler, H., Rinderle-Ma, S.: Process Mining and Conformance Checking of Long Running Processes in the Context of Melanoma Surveillance. International journal of environmental research and public health **15**(12), 2809 (2018)

[34] Rojas, E., Capurro, D.: Characterization of drug use patterns using process mining and temporal abstraction digital phenotyping. In: International Conference on Business Process Management. pp. 187–198. Springer (2018)

[35] Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: A literature review. Journal of biomedical informatics **61**, 224–236 (2016)

[36] 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. Applied Sciences **7**(3) (2017)

[37] Stefanini, A., Aloini, D., Dulmin, R., Mininno, V.: Service Reconfiguration in Healthcare Systems: The Case of a New Focused Hospital Unit. In: Int. Conf. on Health Care Systems Engineering. pp. 179–188. Springer (2017)

[38] Stell, A., Piper, I., Moss, L.: Automated Measurement of Adherence to Traumatic Brain Injury (TBI) Guidelines using Neurological ICU Data. In: International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC). SCITEPRESS (2018)

[39] Tóth, K., Machalik, K., Fogarassy, G., Vathy-Fogarassy, Á.: Applicability of process mining in the exploration of healthcare sequences. In: 30th Neumann Colloquium (NC). pp. 151–156. IEEE (2017)

[40] de Vries, G.J., Neira, R.A.Q., Geleijnse, G., Dixit, P., Mazza, B.F.: Towards Process Mining of EMR Data. In: International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC) (2017)

[41] World Health Organization: International statistical classification of diseases and related health problems, vol. 2. World Health Organization (2004)

[42] Yan, H., Van Gorp, P., Kaymak, U., et al.: Aligning event logs to task-time matrix clinical pathways in BPMN for variance analysis. Journal of biomedical and health informatics **22**(2), 311–317 (2018)

[43] Yang, S., Sarcevic, A., Farneth, R.A., et al.: An approach to automatic process deviation detection in a time-critical clinical process. Journal of biomedical informatics **85**, 155–167 (2018)

[44] Yang, S., Zhou, M., Chen, S., Dong, X., Ahmed, O., Burd, R.S., Marsic, I.: Medical Workflow Modeling Using Alignment-Guided State-Splitting HMM. In: International Conference on Healthcare Informatics (ICHI). pp. 144–153. IEEE (2017)