Process Mining for Case Acquisition in Oncology: A Systematic Literature Review

Joscha Grüger[1], Ralph Bergmann[1,2], Yavuz Kazik[1], and Martin Kuhn[1]

[1] Business Information Systems II, University of Trier, 54286 Trier, Germany
http://www.wi2.uni-trier.de
{grueger,bergmann,s4yakzi,s4makuhn}@uni-trier.de
[2] German Research Center for Artificial Intelligence (DFKI), Branch University of Trier, Behringstraße 21, 54296 Trier, Germany
ralph.bergmann@dfki.de

Abstract Process Mining is a technology family for the analysis of business processes based on event logs. The methods are successfully applied in various areas, including medicine. This paper examines, using a systematic literature review, whether Process Mining is suitable for case acquisition from Hospital Information Systems in order to construct a case base for experience-based systems targeted at decision support in oncology. The review investigates whether there are special characteristics of process mining in the oncological field compared to other medical fields and if the development of similarity measures is discussed in the contributions. For this purpose, 2848 papers were reviewed manually, based on title, abstract and full text, resulting in 55 relevant papers. These were analyzed in detail regarding the research questions. The paper can serve as a basis for further research, identify research opportunities in this domain and provide a useful overview of the current work.

Keywords: Process Mining · Oncology · Case Based Reasoning · literature review.

1 Introduction

Medical guidelines are "systematically developed statements designed to assist healthcare professionals and patients in making decisions about appropriate health care in specific clinical circumstances" [28]. These are classified according to the AWMF[3] system into four development levels from S1 to S3, with S3 being the highest quality level of the development methodology. The classification of a guideline as S3 means that it has undergone all elements of systematic development and the recommendations given therein have a high level of evidence [11]. In the best case, clinicians can make treatment decisions based on these high-quality S3 guidelines and are thus able to offer evidence-based treatment.

Copyright © 2020 by the paper's authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

[3] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften

This is usually possible with well understood disease patterns, such as stroke. In other areas, such as oncology or paediatrics, there is in many cases insufficient evidence for a fully evidence-based treatment of patients. This is partly because studies in these areas are difficult (e.g. with children), diseases are rare or disease patterns are not yet sufficiently researched due to their complexity (e.g. uveal melanoma). In addition, the process of developing guidelines is quite slow, i.e., it usually takes at least two years. In view of scientific progress, especially in medicine, the question of the timeliness of guidelines arises.

In the absence of appropriate guidelines and high evidence studies, treatment decisions are made based on personal experience of medical experts. In contrast to evidence-based medicine, we then speak of "eminence-based medicine", as a treatment decision is based on the comprehensive professional experience of recognized medical experts in the field [19]. In the field of oncology, for example, multidisciplinary experts regularly meet in tumor boards to discuss critical cases and then make decisions, often based on treatment experience with similar patients.

Today, the complexity of such decisions is constantly increasing. The decision-making process is becoming more and more complicated due to the constant development of new therapeutic approaches, an ever-wider range of drugs and their frequently unexplored interaction with given constraints such as comorbidities. In addition, the departure of experienced physicians can have a negative impact on the quality of treatment, as their experience also leaves the clinic.

During the daily treatment of patients, however, physicians systematically record experiential knowledge in hospital information systems (HIS). A HIS is the central information system of a hospital and receives, transmits, processes, stores, and presents information. Date and time of treatments, patient demographics, and examination results are stored in a HIS along with other information [16]. We envision that this information can be used as experience by a Case-Based Reasoning (CBR) system to support eminence-based decision making by the wealth of collected experience available in HIS. For this purpose, treatment processes from a HIS must be captured as a time series of semantically described activities and transferred into semantic case descriptions in order to construct a case base.

In this paper, we therefore investigate based on a literature survey whether process mining, which is an established technology for extracting process knowledge from events logs, can be applied or has been applied already in order to acquire semantic case descriptions from HIS. So far there are only a few literature reviews in the field of process mining in medicine [35,41,13] and only one systematic literature review in the field of process mining in oncology [22]. None of the papers examines the use of process mining for case acquisition for CBR. Processes in the health care sector differ greatly from processes from other domains due to their high complexity, heterogeneity and significant variation over time [17]. This makes it difficult to adapt approaches from other domains. In the present work, a literature study in the medical domain of oncology is performed and used to investigate whether it is possible to generate systematic case descriptions from HIS data using process mining. The paper focuses particularly on the data source from which data is acquired, the process mining methods used, and the data formats and descriptions used, with the aim to provide systematic basis for the topic. By analysing the literature on process mining in oncology, this paper also provides a foundation for future work and helps identifying challenges and research gaps based on the previous research.

The remainder of this paper is organized as follows: in Section 2 we give an overview of the basics of Process Mining and Case Based Reasoning and discuss related work. In Section 3 we present the methodology of the literature review. Then we evaluate the results of the study in Section 4 and summarize them in Section 5 and discuss possible directions for future work.

2 Foundations and Related Work

Case-Based Reasoning [21,3] is an established problem-solving methodology for solving problems based on past experience. Experience is formalized in the form of cases collected in a cases base. A problem (e.g. to determine the best treatment option of a patient) is solved by searching for similar cases in the case base and then reusing the solution contained in the most similar case(s). Unlike black box algorithms such as *deep learning*, the solutions of CBR systems can be easily justified on the basis of similar cases, which can help to strengthen the confidence of healthcare professionals in the AI system, especially in the medical field [26]. The CBR cycle consists of four sequential phases. In the RETRIEVE phase, the most similar cases for a given case are searched for in the case base. Then, in the REUSE phase, the information and knowledge about the most similar cases is used to solve the problem given. Afterwards the solution found in the REVISE phase has to be checked. In the RETAIN phase, those parts of the solution are included in the case base that could be useful for solving later cases [1]. CBR publications in the medical field usually focus exclusively on retrieve and avoid automatic adaptation [8].

Process mining technologies enable the extraction of process knowledge from event logs of information systems. Based on these techniques, process models can be created (discover) and improved (enhancement) and traces can be validated for their conformity with existing models (conformance checking) [37]. Process Mining is already partially used in medicine. The research focuses in particular on the field of oncology and operations. In other areas, such as care giving, cardiology, diabetes, dentistry, medication, intensive care, and radiotherapy, there are considerably fewer publications [13,35]. The focus of most process mining publications in the medical domain is usually on the control flow perspective, based on the discovery of the execution sequence of process activities [35].

3 Methodology

To answer the following research questions, a systematic literature review in the field of process mining in oncology was conducted:

RQ1: What is the state of research in the field of process mining in the domain of oncology?
RQ2: Are there process mining approaches based on oncological data from a HIS?
RQ3: Are there approaches to use process mining for case acquisition for experience-based systems?
RQ4: Are there studies that deal with the similarity of oncological processes?

The search is divided into three main parts: the initial search, the backward snowballing and the forward snowballing [42]. The results of each step are filtered through a three-step application of including- and excluding criteria's (see Fig. 1). Overall, one including, and three excluding criteria were established and

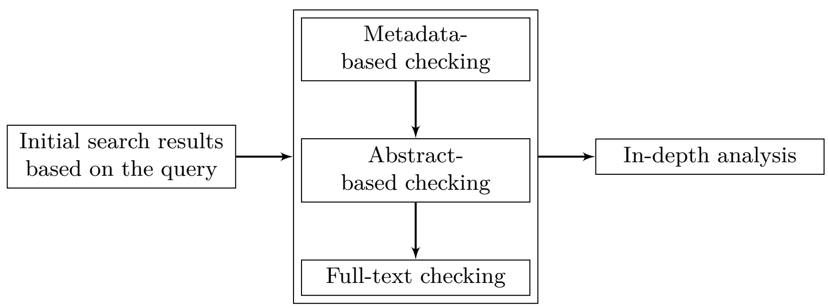

Fig. 1. Applying the including and excluding criteria's.

applied. These ensure that only relevant and accessible documents are included in the analysis:

EC1: Duplicates of the same study are excluded.
EC2: Articles that are not written in English or German are excluded.
EC3: Articles that are not published in a journal or at a conference are excluded.
IC1: Articles written in the field of process mining in oncology or whose authors use oncological data are included.

The first step of the initial search is the database selection. For this purpose, published literature searches in the field of process mining in medicine [35,22,25] were analyzed and the databases used therein were extracted as a basis for database selection. The following sources were identified: ACM DL, CiteSeerX, dblp, Google Scholar, IEEE Explore, PubMed, Science Direct, Scopus, Semantic Scholar, Springer and Web of Science. Based on the databases and a database selection matrix according to Bethel [4,27] the databases Google Scholar and Science Direct were selected.

The search query was created based on the PICOC method (Population, Intervention, Comparison, Outcome, Context) according to Kitchenham [20]. This approach is intended to ensure that the query is precise and only considers the essential components. To ensure that the approach fits the given research question, the Data field has been added and the Comparison and Outcome fields have been removed. The final query is: ("oncology") AND ("process mining") AND ("hospital") AND ("event log"). The same query was used for both databases.

The initial search took place on 20.12.2019. Google Scholar delivered 174 results and Science Direct 24. After forward and backward snowballing, 60 papers were classified as relevant. After analyzing the papers, five papers were excluded due to a lack of information concerning our research questions. Therefore, 55 papers were considered in the analysis process (see Fig. 2).

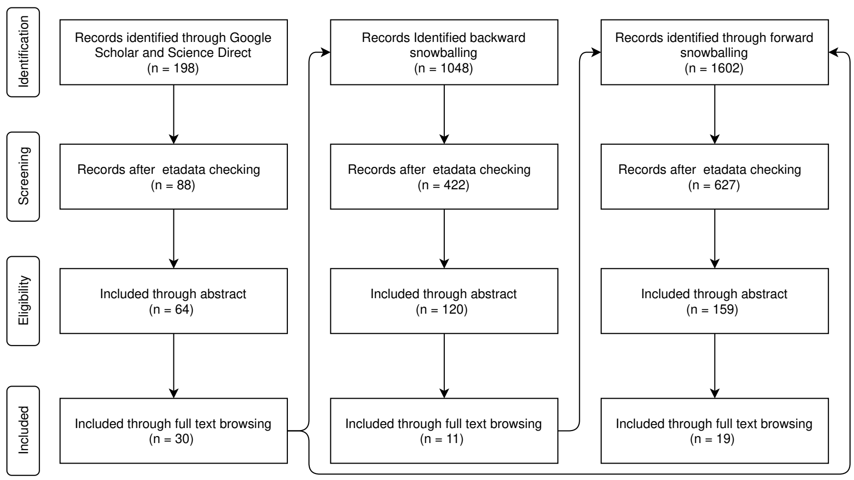

Fig. 2. PRISMA Overview of the results of the different steps of the literature search of the literature review.

To answer the research questions, a data extraction form was developed based on the core features of process mining in oncology and on metadata of the papers.

4 Results

The first papers on process mining in oncology were published in 2008. However, the majority of the papers, 48 out of 55, were published between 2013 and 2019. Most of the papers come from Europe (40 out of 55 papers). With 24 papers the Netherlands is the most important contributor in Europe. This is probably due to the large research group in the field of Process Mining at the University of Eindhoven (TU/e), which was headed by Prof. van der Aalst. From North America and Asia six papers each were found, from South America only two were found.

The papers analyzed address a total of 21 different types of cancer. The majority of the papers referred to gynaecological cancer (19 papers). Other cancers addressed are lung cancer and breast cancer (10 papers each), followed by colorectal cancer (9 papers found), skin cancer (5 papers) and stomach cancer (4 papers). 13 types of cancer were mentioned only once, and in eight contributions the type of tumour was not mentioned.

4.1 Data and Process Mining Perspectives

In order to answer research question RQ2, it was examined on which data the papers work and which data sources were used. After examining the process mining data spectrum, the data used mainly comes from administrative systems (58 %) and from the clinical part of hospital information systems (30 %). Only one paper uses data from medical devices. Most papers, 49 of 55, apply process mining technologies to medical data (diagnosis, prognosis, treatment and prevention of disease activities) and 3 papers use organizational data (management and financial), 3 papers use both medical and organizational data.

Data coming from HIS is described to be very complex, containing heterogeneous structured and unstructured data [10] and sometimes scattered across multiple HIS [5]. Poor data quality and the distribution of data across different HIS can significantly hinder the process extraction [5]. Mans et. al. [36] evaluate data quality issues in the data of a HIS. Among other things, they point out that manual documentation of events leads to the fact that individual events are not documented ("missing events"). In addition, the distribution of the data to different systems leads to imprecise timestamps and executing actors are imprecisely documented (imprecise resource).

Many authors emphasize the complexity of clinical processes (30 papers). They attribute this, among other things, to the high degree of flexibility, the dynamics in treatment processes and in everyday clinical life and a high number of interactions of interdisciplinary actors in a treatment path.

Regarding the process mining perspectives, it can be said that most papers focus on the control flow perspective (48 %). With 23 % follows the time perspective, which was mostly used to identify bottlenecks. The case perspective was only used in 17 % of the papers and the organizational perspective in 11 % of the papers.

The most used process mining technique is process discovery (found in 48 papers). One reason for this is that the other three process mining techniques require a process model, which is often generated via process discovery. Conformance checking was applied in 13 papers and process re-engineering in 6 papers. Operational Support was only used in three papers.

4.2 Process Mining Methodology

The methodology used in the papers clusters the papers according to the tasks to be performed when applying algorithms and techniques for process evaluation. Following [35], the present paper distinguishes between three methodological approaches. The non-domain-specific ad hoc method is used in 21 papers. The clustering method, consisting of the five phases log preparation; log inspection; control flow analysis; performance analysis; and role analysis [6], is used in two papers. The L* life cycle[37], as the third methodological approach, also consists of 5 phases: Planning and justification; extraction; generating the control flow model and linking the event log; generating the integrated process model; and providing operational support [37]. This method was used in 4 papers. Most papers (29 contributions) do not describe a concrete procedure based on known methods.

4.3 Techniques, Algorithms, Tools and Software

In 30 papers special process mining algorithms are used, 36 % of the papers use data mining and machine learning algorithms and 9 % use algorithms from other areas. The most used algorithms are the process discovery algorithms [40] (10 papers), followed by the fuzzy miner [15] (6 papers).

Nearly half of the papers examined use the ProM[4] software (42 %), 7 papers use the R programming language and the Process Mining Toolkit Disco [14] is used in 6 papers. Eight papers have not mentioned any software. ProM is probably the most used tool as it comes with many plugins, offers an interface to develop own plugins and the ProM core is open source[5] [9].

4.4 Clinical Path Similarity

To answer research question RQ4, it was examined which papers cover the similarity of paths. Eight papers deal with the similarity of mined clinical pathways. The main challenge in the application of process mining techniques to medical processes and the subsequent comparison of clinical paths is, in the eyes of 5 out of 8 authors, the flexibility with which the activities are performed. Therefore, many clinical events occur randomly and often without a specified order. Thus, many common similarity measures for processes cannot be applied. Furthermore, it is stated that clinical processes are always time-linked. Therefore, they can change significantly over time and as research progresses [18].

To be able to compare these flexible and heterogeneous clinical pathways, the authors developed and used clustering approaches. The authors used these approaches to cluster activities and then calculated the similarity of the pathways based on the identified clusters of a pathway instead of the specific pathway with treatment activities. Only one approach defines a multidimensional similarity measure and includes besides the pure procedural data also performing actors/resources, and data values to calculate the similarity.

[4] promtools.org

[5] ProM 6 core, GNU Public License

4.5 Process Representation

None of the papers examines explicitly the use of process mining for case acquisition for CBR. Most papers use a procedural process modeling language like Petri Nets [32] (9 papers), BPMN[6] (2 papers) and PWF[7] [12] (2 papers). However, in most cases the exact representation is not given and the procedural character of the process modeling language can only be inferred from the algorithms used. Another representation was chosen by 7 authors, by using a declarative approach. All seven papers chose the declarative process modeling language [38], based on Linear Temporal Logic (LTL). The frequent use of Declare is due to its integration into ProM. The authors usually justify this approach by the suitability of declarative approaches for very flexible processes.

4.6 Research Gaps

To answer research question RQ3, research gaps were identified based on the papers analyzed. For this purpose, the three-step procedure proposed by Müller-Bloch et. al. [31] for identifying research gaps and the PICOS framework [34] was used. This process consists of the localization and characterization of the gaps in step one, the verification of the gaps in step two and the presentation of these in step three. The following research gaps were identified.

No papers were found in the area of case acquisition using process mining for knowledge-based systems (including CBR) in oncology. Studies on the transferability of process mining-based approaches to case acquisition from other domains to oncology are still missing.

One of the papers explicitly examines data quality issues in the process mining context in data from a Dutch hospital. There is no equivalent study for German oncology clinics. The complexity of the data from HIS is mentioned in the papers, but not examined in detail. However, this is interesting for the more advanced and especially for the multi-perspective process mining approaches. Therefore, further studies could provide a basis for further research in this area.

The cancer best researched with process mining technologies is gynecological cancer due to the BPI Challenge data set. Other data sets, such as the MIMIC III data set or the data sets used in [30,23] are not suitable for performance analysis due to data problems [24]. This indicates the urgent need for other available data sources in this domain.

The next gap describes the need of a data quality indicator [2,5,39]. There should be a method to measure the data quality of event logs. This is necessary for unsupervised learning techniques like Sched-Miner which rely on data quality due to the use of unsupervised learning [2]. The three noise types mentioned in [39] are a good starting point for further research concerning the quality indicators.

Research gaps were also identified in process reengineering. Declarative Process Mining deals well with highly variable processes which are the standard for

[6] Business Process Model Notation, https://www.omg.org/spec/BPMN

[7] Pseudo-WorkFlow Language

healthcare processes. In particular, there is a need for research in the preparation of a correct declarative constraint set based on guidelines and an adapted real log to be replayed [33,29].

5 Conclusion and Future Work

The analysis of the papers shows that most papers focus on the analysis of data using process mining and less on describing the process and difficulty of exporting and extracting HIS-data and transforming them into event logs. Data from HIS is described as noisy, incomplete, and complex. This results in a complexity of the mining models, which is due to the lack of data quality on the one hand, but also to the high flexibility of the treatment processes in hospitals.

With regard to process representation and semantification, it can be noted that none of the papers examines the use of process mining for case acquisition for CBR. Most approaches rely on a procedural process modeling language, while 7 papers chose a declarative approach. The authors usually justify this approach by the suitability of declarative approaches for very flexible processes.

In applying similarity measures to oncological processes, the authors see particular challenges in the fact that the processes are highly flexible and change over time as research progresses. Specific challenges for oncological data that differ from other medical domains were not mentioned.

The application of process mining in oncology especially focuses on the control flow perspective. This is probably partly due to the fact that the control flow perspective is often used as the basis for the other process mining perspectives [13]. In terms of methodology, the ad hoc approach is followed mostly by the papers. Compared to the other methodology, it can cope with the complexity of real-world clinical processes [7]. In technical terms, the authors used the heuristic miner most often, arguing that the miner is particularly good at handling noisy data. The most widely used software is ProM.

The results provide a basis for future research in the field of case acquisition from oncological procedural data in HIS using process mining. The investigation of approaches to case acquisition using process mining and the answering of the question of the transferability of the approaches to oncology would be of particular interest. Also, the analysis of data and data quality in German oncology departments in the context of process mining would be of interest for further research. It would also be interesting to systematically investigate the potentials of process mining in CBR approaches.

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI Communications **7**(1), 39–59 (1994). https://doi.org/10.3233/AIC-1994-7104

2. Arik Senderovich, Kyle E. C. Booth, J. Christopher Beck: Learning scheduling models from event data. Proceedings of the International Conference on Automated Planning and Scheduling **29**, 401–409 (2019), https://www.aaai.org/ojs/index.php/ICAPS/article/view/3504
3. Bergmann, R.: Experience Management: Foundations, Development Methodology, and Internet-Based Applications, Lecture Notes in Artificial Intelligence, vol. 2432. Springer (2002)
4. Bethel, A., Rogers, M.: A checklist to assess database-hosting platforms for designing and running searches for systematic reviews. Health information and libraries journal **31**(1), 43–53 (2014). https://doi.org/10.1111/hir.12054
5. Bettencourt-Silva, J.H., Clark, J., Cooper, C.S., Mills, R., Rayward-Smith, V.J., de La Iglesia, B.: Building data-driven pathways from routinely collected hospital data: A case study on prostate cancer. JMIR medical informatics **3**(3), e26 (2015). https://doi.org/10.2196/medinform.4221
6. Caron, F., Vanthienen, J., Baesens, B.: Healthcare analytics: Examining the diagnosis–treatment cycle. Procedia Technology **9**, 996–1004 (2013). https://doi.org/10.1016/j.protcy.2013.12.111
7. Caron, F., Vanthienen, J., Vanhaecht, K., van Limbergen, E., Deweerdt, J., Baesens, B.: A process mining-based investigation of adverse events in care processes. Health information management : journal of the Health Information Management Association of Australia **43**(1), 16–25 (2014). https://doi.org/10.1177/183335831404300103
8. Choudhury, N., Ara, S.: A survey on case-based reasoning in medicine. International Journal of Advanced Computer Science and Applications **7**(8) (2016). https://doi.org/10.14569/IJACSA.2016.070820
9. Claes, J., Poels, G.: Process mining and the prom framework: An exploratory survey. In: La Rosa, M. (ed.) Business Process Management Workshops, Lecture Notes in Business Information Processing, vol. 132, pp. 187–198. Springer Berlin Heidelberg, Berlin/Heidelberg (2013). https://doi.org/10.1007/978-3-642-36285-9_19
10. 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). https://doi.org/10.1016/j.jbi.2016.12.012
11. Encke, A., Kopp, I., Selbmann, H.K.: Bedeutung der S1-, S2-, S3-Leitlinien. Allgemein- und Viszeralchirurgie up2date **3**(04), 257–267 (2009). https://doi.org/10.1055/s-0029-1185952
12. Gatta, R., Lenkowicz, J., Vallati, M., Rojas, E., Damiani, A., Sacchi, L., de Bari, B., Dagliati, A., Fernandez-Llatas, C., Montesi, M., Marchetti, A., Castellano, M., Valentini, V.: pminer: An innovative r library for performing process mining in medicine. In: ten Teije, A., Popow, C., Holmes, J.H., Sacchi, L. (eds.) Artificial intelligence in medicine, Lecture notes in computer science Lecture notes in artificial intelligence, vol. 10259, pp. 351–355. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59758-4_42
13. Ghasemi, M., Amyot, D.: Process mining in healthcare: a systematised literature review. International Journal of Electronic Healthcare **9**(1), 60 (2016). https://doi.org/10.1504/IJEH.2016.078745
14. Günther, C.W., Rozinat, A.: Disco: discover your processes. In: Lohmann, N., Moser, S. (eds.) Proceedings of the Demonstration Track of the 10th International Conference on Business Process Management (BPM 2012). pp. 40–44. CEUR Workshop Proceedings, CEUR-WS.org (2012)

15. Günther, C.W., van der Aalst, W.M.P.: Fuzzy mining – adaptive process simplification based on multi-perspective metrics. In: Alonso, G., Dadam, P., Rosemann, M. (eds.) Business process management, Lecture Notes in Computer Science, vol. 4714, pp. 328–343. Springer, Berlin (2007). https://doi.org/10.1007/978-3-540-75183-0_24
16. Haux, R.: Strategic information management in hospitals: An introduction to hospital information systems. Health informatics series, Springer, New York (2004), http://www.loc.gov/catdir/enhancements/fy0818/2003059129-d.html
17. Homayounfar, P.: Process mining challenges in hospital information systems. pp. 1135–1140 (01 2012)
18. Huang, Z., Gan, C., Lu, X., Huan, H.: Mining the changes of medical behaviors for clinical pathways. Studies in health technology and informatics **192**, 117–121 (2013)
19. Kelly, A.M.: Evidence-based practice: an introduction and overview. Seminars in roentgenology **44**(3), 131–139 (2009). https://doi.org/10.1053/j.ro.2009.03.010
20. Kitchenham, B., Charters, S.: Guidelines for performing systematic literature reviews in software engineering (2007)
21. Kolodneer, J.L.: Improving human decision making through case-based decision aiding. AI Magazine **12**(2), 52 (1991). https://doi.org/10.1609/aimag.v12i2.895, https://www.aaai.org/ojs/index.php/aimagazine/article/view/895
22. Kurniati, A.P., Johnson, O., Hogg, D., Hall, G.: Process mining in oncology: A literature review. In: Proceedings of the 6th International Conference on Information Communication and Management ICICM 2016. pp. 291–297. IEEE Press, Piscataway, NJ (2016). https://doi.org/10.1109/INFOCOMAN.2016.7784260
23. Kurniati, A.P., McInerney, C., Zucker, K., Hall, G., Hogg, D., Johnson, O.: A multi-level approach for identifying process change in cancer pathways. In: Di Francescomarino, C., Dijkman, R., Zdun, U. (eds.) BUSINESS PROCESS MANAGEMENT WORKSHOPS, Lecture Notes in Business Information Processing, vol. 362, pp. 595–607. Springer, [Place of publication not identified] (2020). https://doi.org/10.1007/978-3-030-37453-2_48
24. Kurniati, A.P., Rojas, E., Hogg, D., Hall, G., Johnson, O.A.: 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 **25**(4), 1878–1893 (2019). https://doi.org/10.1177/1460458218810760
25. Kusuma, G.P., Hall, M., Gale, C.P., Johnson, O.A.: Process mining in cardiology: A literature review. International Journal of Bioscience, Biochemistry and Bioinformatics **8**(4), 226–236 (2018). https://doi.org/10.17706/ijbbb.2018.8.4.226-236
26. Lamy, J.B., Sekar, B., Guezennec, G., Bouaud, J., Séroussi, B.: Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach. Artificial Intelligence in Medicine **94**, 42–53 (2019). https://doi.org/10.1016/j.artmed.2019.01.001
27. Levay, P., Craven, J.: Systematic searching: Practical ideas for improving results. Facet Publishing (2019)
28. Lohr, K.N., Field, M.J. (eds.): Clinical practice guidelines: Directions for a new program, Publication IOM, vol. 90-08. National Academy Press, Washington, D.C (1990). https://doi.org/10.17226/1626
29. Maggi, F.M., Bose, R.P.J.C., van der Aalst, W.M.P.: A knowledge-based integrated approach for discovering and repairing declare maps. In: Salinesi, C., Norrie, M.C., Pastor, O. (eds.) Advanced Information Systems Engineering, Lecture Notes in Computer Science, vol. 7908, pp. 433–448. Springer Berlin Heidelberg, Berlin/Heidelberg (2013). https://doi.org/10.1007/978-3-642-38709-8_28

30. Meng, W., Ou, W., Chandwani, S., Chen, X., Black, W., Cai, Z.: Temporal phenotyping by mining healthcare data to derive lines of therapy for cancer. Journal of Biomedical Informatics **100**, 103335 (2019). https://doi.org/10.1016/j.jbi.2019.103335
31. Müller-Bloch, C., Kranz, J.: A framework for rigorously identifying research gaps in qualitative literature reviews. In: ICIS (2015)
32. Peterson, J.L.: Petri nets. ACM Computing Surveys (CSUR) **9**(3), 223–252 (1977). https://doi.org/10.1145/356698.356702
33. Rinner, C., Helm, E., Dunkl, R., Kittler, H., Rinderle-Ma, S.: An application of process mining in the context of melanoma surveillance using time boxing. In: Daniel, F., Sheng, Q.Z., Motahari, H. (eds.) Business Process Management Workshops, Lecture Notes in Business Information Processing, vol. 342, pp. 175–186. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-11641-5_14
34. Robinson, K.A., Saldanha, I.J., Mckoy, N.A.: Frameworks for Determining Research Gaps During Systematic Reviews. Rockville (MD) (2011)
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). https://doi.org/10.1016/j.jbi.2016.04.007
36. Rs Ronny Mans, Van der Aalst, Rjb Rob Vanwersch: Process mining in healthcare : opportunities beyond the ordinary. Computer Science (2013)
37. van der Aalst: Process mining manifesto. In: Daniel, F., Barkaoui, K., Dustdar, S. (eds.) Business Process Management Workshops, Lecture Notes in Business Information Processing, vol. 99, pp. 169–194. Springer Berlin Heidelberg, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28108-2_19
38. van der Aalst, W.M.P., Pesic, M., Schonenberg, H.: Declarative workflows: Balancing between flexibility and support. Computer Science - Research and Development **23**(2), 99–113 (2009). https://doi.org/10.1007/s00450-009-0057-9
39. van der Spoel, S., van Keulen, M., Amrit, C.: Process prediction in noisy data sets: A case study in a dutch hospital. In: Mylopoulos, J., Rosemann, M. (eds.) Data-Driven Process Discovery and Analysis, Lecture Notes in Business Information Processing, vol. 162, pp. 60–83. Springer Berlin Heidelberg, Berlin/Heidelberg (2013). https://doi.org/10.1007/978-3-642-40919-6_4
40. Weijters, A.J.M.M., Aalst, van der, W.M.P., Alves De Medeiros, A.K.: Process mining with the HeuristicsMiner algorithm. BETA publicatie : working papers, Technische Universiteit Eindhoven (2006)
41. Williams, R., Rojas, E., Peek, N., Johnson, O.A.: Process mining in primary care: A literature review. Studies in health technology and informatics **247**, 376–380 (2018)
42. Wohlin, C.: Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: Shepperd, M., Hall, T., Myrtveit, I. (eds.) Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering. pp. 1–10. ACM, New York, NY (2014). https://doi.org/10.1145/2601248.2601268