ISSN: 0976-3104
SPECIAL ISSUE: Healthcare Systems Management

ARTICLE

# PROCESS MINING FOR CHECK-UP PROCESS ANALYSIS

Onur Dogan*

*Department of Industrial Engineering, Istanbul Technical University, Istanbul, TURKEY*

## ABSTRACT

***Background:*** *Healthcare environment has a critical importance for several reasons. First of all, health is one of the most valuable wealth in human life. There have been many studies to develop clinical processes by optimizing resources, reducing the waiting time, increasing the patient satisfactory. In Turkey, majority of the healthcare processes are monitored, managed and improved in ad hoc, manual and labor-intensive ways. Since healthcare processes are complicated and flexible, and have huge number of data, this process management methods are insufficient and ineffective.* ***Methods:*** *In this study, we use process mining technique to analyze check-up processes. Process mining uses event logs to automatically extract information about the process. It can be seen as a missing link between data science and process science.* ***Results:*** *We examined 372 events, 732 cases, and 9 different activities in this study. Bloodletting is the mostly occurred activity. At the same time, it increases 'Patient Record' occurrence due to the dominant loop. The average duration of a check-up is 59.5 hours.* ***Conclusions:*** *Process mining techniques provides an easy to use way to success a global view of the processes. With the results of the process mining, the health professionals and managers can achieve a real view of the problems that are currently happening in the analyzed area. It also helps to improve complex processes in hospitals.*

**KEY WORDS**
*process mining, checkup, workflow, event log, patient paths*

Received: 20 June 2018
Accepted: 30 July 2018
Published: 15 Sept 2018

***Corresponding Author**
Email: odogan@itu.edu.tr
Tel.: +90 212 293 13 00
Fax: +90 212 240 72 60

## INTRODUCTION

Process mining has been applied in many areas such as education [1, 2], banking [3, 4], manufacturing [5], municipality [6, 7], and informatics [8, 9]. These domains have often partly structured or semi-structured processes with some exceptional behaviors. However, healthcare environments need flexible decision making and have mainly unstructured processes. Explicitly, a healthcare center is not a factory and patients are more valuable than products. Healthcare domains have a critical significance for several reasons. Health is one of the most valuable wealth in human life. Demand of medical services has increased due to aging population and improved standards of living. The number of cancer patients in Turkey has increased rapidly. Check-up as a precaution reaction has a critical significance to avoid cancer types and other types of diseases.

Check up is a kind of health screening taking into account age, hereditary structure and environmental factors. The aim of the check-up is to determine the possible diseases of a person without health problems in the early period and to take precautions. Health problems that are difficult to treat in advanced stages; it leads to a depressing process for the patient and his family. Early diagnosis and planned treatment can be life-saving, especially after a yearly health screening in people with a family history.

To optimize a process, one has to first understand the As-Is process. And this is usually far from simple, because business processes are performed by a number of people, often across different organizational units or even companies. Everybody only sees a part of the process. In majority of industrial applications, ad hoc, manual and labor-intensive ways were used to improve, monitor and analyze a healthcare process [10]. However, these methods have some limitations. The effectiveness and success of an ad hoc approach can only be measured by checking the impact on performance metrics. This impact analysis needs several months, which is not reasonable for healthcare process management. Process variations and exceptions and their root causes are critical factors that affect the process performance indicators. Capturing these factors with a tradition process management methods is not easy. Moreover, almost all practices have differences between prescribed and actual processes.

Process mining gains insight into actual healthcare processes. It can capture process variations and process exceptions. The root causes of an ineffectiveness event can be defined. In this study, we propose the application of process mining to analyze check-up processes in an oncology hospital in Turkey. The data for over 630 patient activities, extracted from the hospital information systems, were linked together and analysed to better understand the differences in the practices associated with check-up management.

## RELATED WORKS

Process mining many advantages for professionals in different sectors. Process mining implementations have grown in especially healthcare environment because most healthcare processes are complex, dynamic, and multi-disciplinary. According to the study of Erdogan and Tarhan [11], process mining has rapidly grown in healthcare area.

Diagnosis and treatment processes in hospitals usually vary for each patient. The reason for the variability may be an effective tool to improve processes. Caron, Vanthienen, and Baesens [12] investigated clinical processes using past event logs to uncover variabilities. They aimed to produce positive outputs such as optimum resource usage and increasing patient satisfactory and safety. It is very important to understand real processes to reduce costs and improve quality. Montani et al. [13, 14] applied case retrieve and

Guest Editor: Prof. Dr. Haticie Camgöz Akdag

IIOAB JOURNAL ISSN: 0976-3104

process mining techniques to stroke processes. Lin et al. [15] examined the processes of cerebral palsy patients with process mining and predicted the path of new patients with data mining methods. Yoo et al. [16] analyzed the outpatient processes in varying conditions such as environmental changes, waiting time, and the time spent for treatment. Rojas et al. [17] focused on the application of process mining and data analysis techniques to answer questions about emergency room processes that are frequently asked by emergency room experts. Badakhshan and Alibabaei [18] used three types of process mining which are discovery, conformance checking and enhancement, in emergency call services. Rebuge and Ferreira [19] developed a methodology for process mining techniques, which identified regular behavior, process variables, and exceptional medical conditions. They used sequence clustering method to compile the behavior of successive traces in the logs. The complexity of the processes makes the value stream mapping difficult in hospitals. Antonelli and Bruno [20] used process mining to draw value stream mapping. Partington et al. [21] applied process mining to measure differences in the treatment of patients with breast pain. They focused on inter-organization comparison of processes and process performances. To facilitate comparison analysis, they implemented clustering, process discovery, performance analysis and flows. Blum et al. [22] proposed a model for surgeries to automatically visualize and produce a statistical model describing the operation flow. Since the management and effectiveness of clinical processes were not studied, Overduin [23] filled this gap. He made an application in the process of cataract surgery.

In a general manner, check-up is performed for several purposes. The check-up type may change according to the patient needs. Several types such as full or partial may be used depending on the patient health condition. The main aim of the study is to discover the most followed patient paths and uncover deviations from prescribed hand-made process model.

## PROCESS MINING

Process mining is a methodology that covers not only process modeling and process analysis abut also data mining and business intelligence [24]. The major goal of process mining is to discover, explore, control and develop business processes from event logs in the information systems. Process mining covers several study areas such as process discovery, comparison of the model and event log, checking of deviations, social network analysis, simulation models, forecasting and recommendations [25]. In the era of Industry 4.0, the digital world and the physical world are now intertwined. Information systems record huge amount of event logs that can be used to manage processes. Process mining uses the recorded event logs as an input to analyze the business processes.

Discovery, conformance checking, and enhancement are the types of the process mining [25]. The process discovery is the most important and popular output of a process mining study [26]. Process discovery algorithms create a process model without using any prior knowledge. Different notations such as Petri Net and BPMN describe the behavior in the event logs and represent the discovered model. There are several discovery algorithms such as alpha [27], genetic mining (De Medeiros et al., 2007), PALIA (Fernández-Llatas et al., 2010), and heuristic mining [30].

Although we will not go details of the process mining principles in this study, they are necessay to undertand how process mining works. Process mining set log-based ordering relations among activities. Disco scans all event logs. If activity *a* is followed by activity *b*, and activity *b* is never followed by activity *a*, it is assumed that there is a causal dependency between activity *a* and activity *b*.

**Definition 1:** Log-based ordering relations [25]: Let $L$ be an event log over $A$ and let $a,b \in A$.

$a >_L b$; if and only if there is a trace $\sigma = \langle t_1, t_2, t_3, \ldots, t_n \rangle$ and $i \in \{1,\ldots,n-1\}$ such that $\sigma \in L$, $t_i = a$ and $t_{i+1} = b$;

$a \rightarrow_L b$; if and only if $a >_L b$ and $b \not>_L a$;

$a \#_L b$; if and only if $a \not>_L b$ and $b \not>_L a$;

$a \|_L b$; if and only if $a >_L b$ and $b >_L a$;

For example, $L = \left[ \langle k,l,m,n \rangle^3, \langle k,m,l,n \rangle^2, \langle k,o,n \rangle \right]$ represents the traces in the event logs. In the event log, the order of $\langle k,l,m,n \rangle$ is counted three times. For this event log, following log-based ordering relations can be created.

$$>_L = \{(k,l),(k,m),(k,o),(l,m),(m,l),(l,n),(m,n),(o,n)\}$$

$$\rightarrow_L = \{(k,l),(k,m),(k,o),(l,n),(m,n),(o,n)\}$$

$$\#_L = \{(k,k),(k,n),(l,l),(l,o),(m,m),(m,o),(n,k),(n,n),(o,l),(o,m),(o,o)\}$$

$$\|_L = \{(l,m),(m,l)\}$$




Guest Editor: Prof. Dr. Haticie Camgöz Akdag

IIOAB JOURNAL ISSN: 0976-3104

The relation $>_L$ shows all activity pairs in a direct follows. For example, *k* is directly followed by *l* in the trace of $\langle k,l,m,n\rangle$. On the other hand, $m \not>_L n$ because *m* does not directly follow *n* in any trace. The relation $\rightarrow_L$ presents all relation activity pairs that have a causality relation. $m \rightarrow_L n$ because *c* sometimes directly follows *d* and *d* never directly follows *c*. The relation $\#_L$ refers all activity pairs that does not directly follow each other. Since *l* never follows *o* and never the other way around, therefore $l \#_L o$. The relation $\|_L$ includes parallel activity pairs. $l \|_L m$ because sometimes *b* directly follows *c* and sometimes vice versa. In other words, $b >_L c$ and $c >_L b$. A footprint matrix shows relations among activity pairs. [Table 1] represents the footprints for the event log *L*.

**Table 1.** Footprint of event log *L*

| | k | l | m | n | o |
|---|---|---|---|---|---|
| **k** | $\#_L$ | $\rightarrow_L$ | $\rightarrow_L$ | $\#_L$ | $\rightarrow_L$ |
| **l** | $\leftarrow_L$ | $\#_L$ | $\|_L$ | $\rightarrow_L$ | $\#_L$ |
| **m** | $\leftarrow_L$ | $\|_L$ | $\#_L$ | $\rightarrow_L$ | $\#_L$ |
| **n** | $\#_L$ | $\leftarrow_L$ | $\leftarrow_L$ | $\#_L$ | $\leftarrow_L$ |
| **o** | $\leftarrow_L$ | $\#_L$ | $\#_L$ | $\rightarrow_L$ | $\#_L$ |

## CASE STUDY

For this study, we gathered data from a private hospital in Istanbul. A full check-up process has several activities; MR, BT and USG in radiology department, bloodletting, EKG, lung x-ray, abdominal ultrasonography, mammography, other examinations and preparation the check-up report. According to the patients need, the check-up process may vary.

Some basic process mining terms can be explained briefly to better understand the study.
*Activities:* Total number of different activity types in the event log.
*Events:* Total number of activities in the data set.
*Cases:* Total number of process instances in the event log.
*Attributes:* Total number of columns from the dataset that have been imported.
*Start and End:* The range of time covered by event log from earliest to latest timestamp observed.

### Data preparation

The final output in Microsoft Access was a table containing the following columns: (1) customerID, (2) activity (3) start and (4) end. Subsequently the final output from Microsoft Access was exported to a Microsoft Excel spreadsheet. From Microsoft Excel it could be directly imported in the process mining tool Disco. In the Disco import tool, the column ‘customerID’ was selected as the cases, the columns ‘start’ and ‘end’ as timestamps, the column ‘activity’ as events. Table 2 shows a sample of event logs.

**Table 2.** A Sample of the Event Log

| CustomerID | ActivityName | Start | End |
|---|---|---|---|
| 34612045 | Patient Record | 15.12.2017 10:01 | 15.12.2017 10:10 |
| 38541418 | Patient Record | 15.12.2017 10:07 | 15.12.2017 11:17 |
| 39063121 | Patient Record | 15.12.2017 10:09 | 15.12.2017 10:21 |
| 34612045 | Lung X-Ray | 15.12.2017 10:11 | 15.12.2017 10:38 |
| 33109839 | Abdominal Ultrasonography | 15.12.2017 10:12 | 15.12.2017 10:25 |
| 34612045 | Patient Record | 15.12.2017 10:19 | 15.12.2017 10:24 |
| 40271267 | Bloodletting | 15.12.2017 10:21 | 15.12.2017 10:33 |
| 39063121 | Other Examinations | 15.12.2017 10:21 | 15.12.2017 10:42 |
| 34612045 | Patient Record | 15.12.2017 10:01 | 15.12.2017 10:10 |
| 38541418 | Patient Record | 15.12.2017 10:07 | 15.12.2017 10:17 |

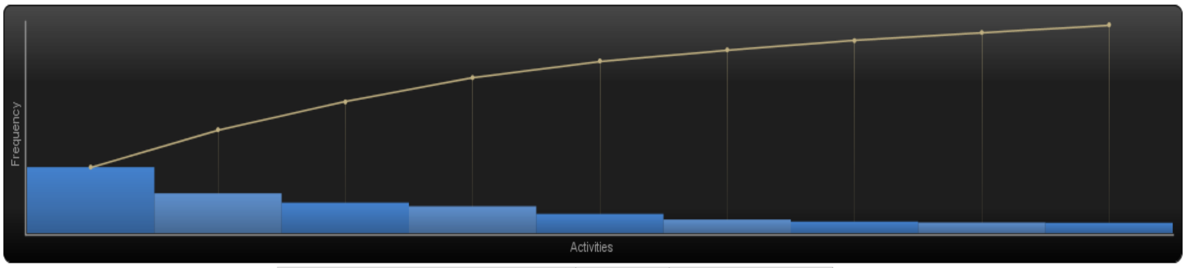
### Discovery of Check-up Process and Evaluation

Log inspection was carried out in Disco. In total, the event-log contained 2372 events, 732 cases, and 9 distinct events (activities). Although basic statistics do not give results on the general view of the patient pathways, they are useful to see some data-centric results. Figure 1 shows the activity details.

THE IIOAB JOURNAL

Guest Editor: Prof. Dr. Haticie Camgöz Akdag

IIOAB JOURNAL
ISSN: 0976-3104




| Activity | Frequency | Relative frequency |
|---|---|---|
| Patient Record | 732 | 30.86% |
| Bloodletting | 430 | 18.13% |
| Abdominal Ultrasonography | 321 | 13.53% |
| MR, BT, USG | 280 | 11.80% |
| Check-up Report | 192 | 8.09% |
| Lung X-Ray | 128 | 5.40% |
| EKG | 105 | 4.43% |
| Other Examinations | 94 | 3.96% |
| Mammography | 90 | 3.79% |

**Fig. 1:** Activity frequencies.

Figure 2 shows the discovered process model in Disco. The process map is the most important analysis result in Disco. It shows you how your process has actually been executed. The process flows that you see in the Map view are automatically reconstructed (”discovered”) based on the sequence and timing of the activities in your imported event log data. So, without further knowledge about the process, or any pre-existing process model, you obtain an objective picture of the real process.

In the discovered process model, only the most dominant paths in the process map is used to show the paths instead up to all connections between activities that have occurred. This means that only the most dominant connections between these activities are shown. Disco makes sure that all your activities are always connected and avoids getting “dangling” process fragments that cannot be put in context with the remaining activities even if you look at a simplified process map. On the other hand, all activities in the process are shown. ‘Bloodletting’ has been performed 200 times directly after the activity ‘Patient Record’ but 176 times the process has returned to ‘Patient Record’. This is because we use only the most dominant paths.

Activities are represented by boxes and the process flow between two activities is visualized by an arrow. For example, there are 430 cases in the data set that all start with the activity Bloodletting. Dashed arrows point to activities that occurred at the very beginning or at the very end of the process. The absolute frequencies are displayed in the numbers at the arcs and in the activities. For instance, after the activity ‘Patient Record’, the process splits into three alternative paths: In 90 cases the activity ‘MR, BT, USG’ was performed, ‘Bloodletting’ was performed in 200 cases. In 203 cases, patients directly left the check-up system. The thickness of the arrows and the coloring of the activities visually support these numbers. Because the path where 200 cases have “travelled through” indicates the main flow in this part of the process, it is visualized by a thicker arrow.

In our study, the activity ‘Patient Record’ means both saving patient details and iformation desk. Therefore, the activity ‘Patient Record’ is normally the one that is executed most often (in total 732 times). This comes from the dominant loop with activity ‘Bloodletting’. Repeatedly, ‘Bloodletting’ activity is amended and need to be re-analyzed, which is of course very inefficient and from a process improvement perspective we would need to find out what is going on. Perhaps patients don’t know how they are performed to examinatation, and we might resolve the problem by updating the check-up guidelines or providing additional training. 192 check-up reports are completed and created a report. Some others in 203 cases, are stopped earlier in the process.

The graph in Figure 3 shows the mean activity duration of the process. There are 2372 events in 630 different cases. The average duration of a check-up is 59.5 hours.

THE IIOAB JOURNAL

**Guest Editor:** Prof. Dr. Haticie Camgöz Akdag

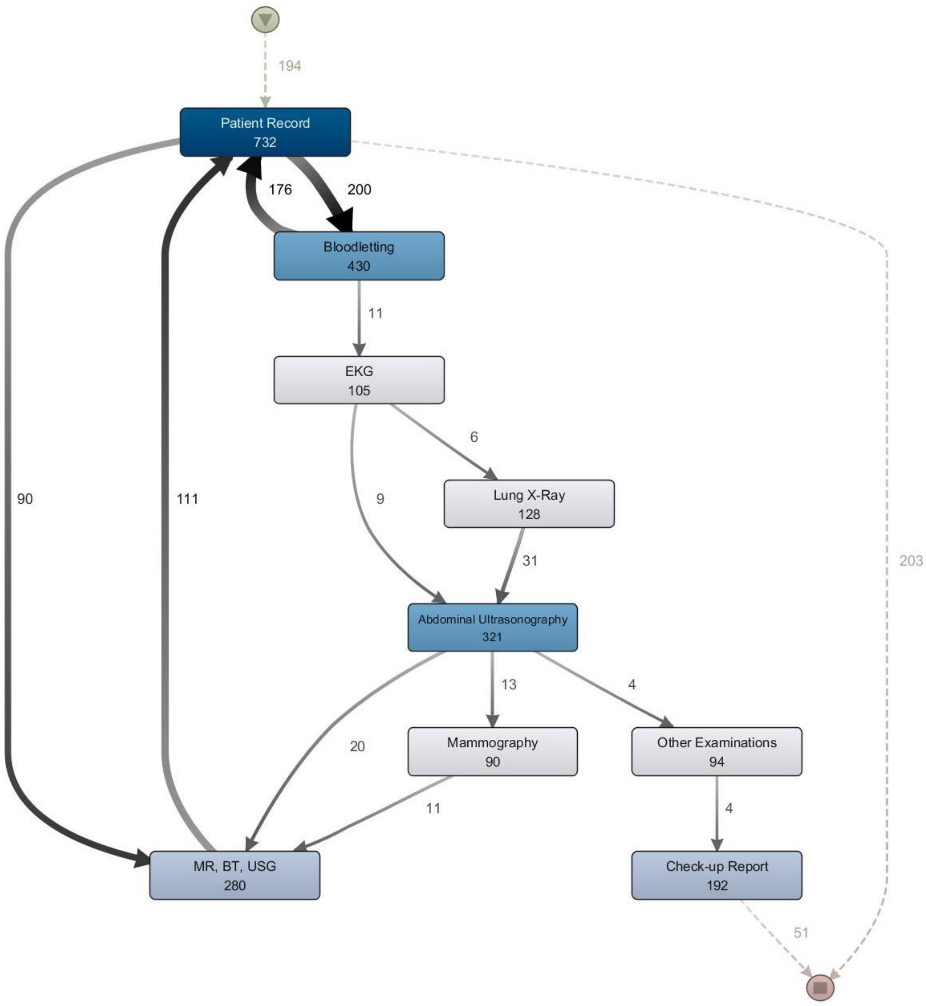



**Fig. 2:** Discovered process model.

| Case ID | Events | Variant | Started | Finished | Duration |
|---|---|---|---|---|---|
| 34612045 | 9 | Variant 54 | 15.12.2017 10:01:00 | 16.12.2017 10:05:00 | 1 day, 4 mins |
| 38541418 | 3 | Variant 10 | 15.12.2017 10:07:00 | 17.12.2017 09:59:00 | 1 day, 23 hours |
| 39063121 | 2 | Variant 37 | 15.12.2017 10:09:00 | 15.12.2017 10:22:00 | 13 mins |
| 33109839 | 9 | Variant 55 | 15.12.2017 10:12:00 | 16.12.2017 12:35:00 | 1 day, 2 hours |
| 40271267 | 7 | Variant 56 | 15.12.2017 10:21:00 | 16.12.2017 17:08:00 | 1 day, 6 hours |
| 35261493 | 29 | Variant 57 | 15.12.2017 10:22:00 | 16.12.2017 10:58:00 | 1 day, 36 mins |
| 33683153 | 24 | Variant 58 | 15.12.2017 10:25:00 | 16.12.2017 10:15:00 | 23 hours, 50 mins |
| 17103096 | 6 | Variant 59 | 15.12.2017 10:29:00 | 28.12.2017 19:14:00 | 13 days, 8 hours |
| 35303834 | 4 | Variant 60 | 15.12.2017 10:29:00 | 16.12.2017 11:31:00 | 1 day, 1 hour |
| 42615791 | 3 | Variant 48 | 15.12.2017 10:30:00 | 15.12.2017 11:54:00 | 1 hour, 24 mins |
| 18015162 | 4 | Variant 61 | 15.12.2017 10:32:00 | 16.12.2017 14:44:00 | 1 day, 4 hours |
| 38428442 | 1 | Variant 1 | 15.12.2017 10:32:00 | 15.12.2017 10:40:00 | 8 mins |
| 41639760 | 13 | Variant 62 | 15.12.2017 10:36:00 | 16.12.2017 10:42:00 | 1 day, 6 mins |
| 16767780 | 3 | Variant 24 | 15.12.2017 10:40:00 | 18.12.2017 10:13:00 | 2 days, 23 hours |
| 38583778 | 1 | Variant 2 | 15.12.2017 10:41:00 | 15.12.2017 11:01:00 | 20 mins |
| 25899048 | 4 | Variant 63 | 15.12.2017 10:43:00 | 18.12.2017 11:38:00 | 3 days, 55 mins |
| 31236245 | 10 | Variant 64 | 15.12.2017 10:43:00 | 23.12.2017 16:26:00 | 8 days, 5 hours |
| 34163297 | 10 | Variant 65 | 15.12.2017 10:45:00 | 16.12.2017 12:42:00 | 1 day, 1 hour |
| 30743518 | 5 | Variant 66 | 15.12.2017 10:48:00 | 18.12.2017 11:06:00 | 3 days, 18 mins |
| 25754902 | 9 | Variant 67 | 15.12.2017 10:50:00 | 19.12.2017 17:40:00 | 4 days, 6 hours |
| 38653230 | 3 | Variant 68 | 15.12.2017 10:52:00 | 27.12.2017 19:20:00 | 12 days, 8 hours |
| 32918407 | 3 | Variant 20 | 15.12.2017 10:55:00 | 18.12.2017 14:59:00 | 3 days, 4 hours |

**Fig. 3:** Mean activity duration.

## CONCLUSION

According to the study, the application of process mining techniques in combination with hospital information systems provides an easy to use way to success a global view of the processes. In this work, we have stated that the process mining can capture the features of the processes, showing them in an easy and understandable view that is accepted by the medical staff in a real environment. With this information, the health professionals and managers can achieve a real view of the problems that are currently happening in the analyzed area. This enables them the improvement of processes with a better



Guest Editor: Prof. Dr. Haticie Camgöz Akdag

knowledge of the problems, increasing their efficiency and the probability of success for their further deployment in the real context.

In order to apply the results achieved in this study to any another context, it is necessary to deal with the spaghetti effect limitation. The spaghetti effect is a well-known effect that decrease the understandability of flows in very complex problems. Using different levels of Path slider and Activity slider, the level of detail in the process model can be adjusted.

For further researches, increasing the filters over the event log can provide interesting views about the process in order to achieve specific knowledge of some details of the process. For example, it is possible to infer the process flow of patients with a specific check-up stage. In this way, this paradigm can be very useful to achieve more and better findings for an increase the quality of service in healthcare centers.

## CONFLICT OF INTEREST

The author declares no conflict of interest

## ACKNOWLEDGEMENTS

None

## FINANCIAL DISCLOSURE

None

## REFERENCES

[1] Schoor C, Bannert M. [2012] Exploring regulatory processes during a computer-supported collaborative learning task using process mining. Comput. Human Behav. 28(4):1321–1331.

[2] Juhaňák L, Zounek J, Rohlíková L. [2017] Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. Comput. Human Behav. 92:496-506.

[3] de Weerdt J, Schupp A, Vanderloock A, Baesens B. [2013] Process Mining for the multi-faceted analysis of business processes - A case study in a financial services organization. Comput. Ind. 64(1):57-67.

[4] Werner M, Gehrke N. [2015] Multilevel Process Mining for Financial Audits. IEEE Trans. Serv. Comput. 8(6):820:832.

[5] Ou-Yang C, Winarjo H. [2011] Petri-net integration - An approach to support multi-agent process mining. Expert Syst. Appl. 38(4):4039–4051.

[6] van der Aalst WMP, Reijers HA, Weijters AJ, van Dongen BF, De Medeiros AA, Song M, Verbeek HMW. [2007] Business process mining: An industrial application. Inf. Syst. 32(5):713–732.

[7] van der Aalst W.M.P., Schonenberg M. H., Song M. [2011] Time prediction based on process mining. Inf. Syst. 36(2):450-475.

[8] Valle AM, Santos EAP, Loures ER. [2017] Applying process mining techniques in software process appraisals. Inf. Softw. Technol. 87: 19–31.

[9] Folino F, Guarascio M, Pontieri L. [2017] A descriptive clustering approach to the analysis of quantitative business-process deviances. Proc. Symp. Appl. Comput. (SAC '17). ACM, New York, NY, USA 765–770.

[10] Poelmans J, Dedene G, Verheyden G, van der Mussele H, Viaene S, Peters E. [2010] Combining business process and data discovery techniques for analyzing and improving integrated care pathways. In Industrial Conference on Data Mining. Springer, Berlin, Heidelberg 505–517.

[11] Erdogan TG, Tarhan A. [2018] Systematic Mapping of Process Mining Studies in Healthcare. IEEE Access 6:24543–25567.

[12] Caron F, Vanthienen J, Baesens B. [2013] Healthcare Analytics: Examining the Diagnosis–treatment Cycle. Procedia Technol. 9:996–1004.

[13] Montani S, Leonardi G, Quaglini S, Cavallini A, Micieli G. [2014] Improving structural medical process comparison by exploiting domain knowledge and mined information. Artif. Intell. Med. 62(1):33–45.

[14] Montani S, Leonardi G, Quaglini S, Cavallini A, Micieli G. [2013] Mining and retrieving medical processes to assess the quality of care. Lect. Notes Comput. Sci. 7969:233–240.

[15] Lin F, Chou S, Pan S, Chen Y. [2001] Mining Time Dependency Patterns in Clinical Pathways. Proc. 33rd Hawaii Int. Conf. Syst. Sci. 62(1):11-25.

[16] Yoo S, Cho M, Kim E, Kim S, Sim Y, Yoo D, Song M. [2016] Assessment of hospital processes using a process mining technique: outpatient process analysis at a tertiary hospital. Int. J. Med. Inform. 88:34-43.

[17] Rojas E, Sepúlveda M, Munoz-Gama J, Capurro D, Traver V, Fernandez-Llatas C. [2017] Question-Driven Methodology for Analyzing Emergency Room Processes Using Process Mining. Appl. Sci. 7(3):302-332.

[18] Badakhshan P, Alibabaei A. [2018] Using Process Mining for Process Analysis Improvement in Pre-Hospital Emergency. In Inf. and Com. Tech. in Org. and Soc. 425–439.

[19] Rebuge A, Ferreira DR. [2012] Business process analysis in healthcare environments: A methodology based on process mining. Inf. Syst. 37(2):99-116.

[20] Antonelli D, Bruno G. [2015] Application of Process Mining and Semantic Structuring Towards a Lean Healthcare Network. Adv. Inf. and Com. Tech. 463:97–508.

[21] Partington A, Wynn MT, Suriadi S, Ouyang C, Karnon J. [2015] Process mining for clinical processes: A comparative analysis of four Australian hospitals. ACM Trans. Manag. Inf. Syst. 5(4):339-350.

[22] Blum T, Padoy N, Feußner H, Navab N. [2008] Workflow mining for visualization and analysis of surgeries. Int. J. Comput. Assist. Radiol. Surg. 3(5):379-386.

[23] Overduin [2013] Exploration of the link between the execution of a clinical process and its effectiveness using process mining techniques. Master Theses in School of Industrial Engineering at TU/e.

[24] van der Aalst WMP. [2017] Decomposing Petri Nets for Process Mining: A Generic Approach. Distrib. Parallel Databases 31(4):471-507.

[25] van der Aalst WMP. [2011] Process Mining: Data Science in Action. Second Edi. Springer.

[26] van der Aalst WMP. [2007] Challenges in Business Process Analysis. in International Conference on Enterprise Information Systems 27–42.

[27] van der Aalst WMP, van Dongen BF, Herbst J, Maruster L, Schimm G, Weijters AJMM. [2003] Workflow mining: A survey of issues and approaches. Data Knowl. Eng. 47(2):237-267.

[28] de Medeiros AKA, Weijters AJMM, van der Aalst WMP. [2007] Genetic process mining: An experimental evaluation. Data Min. Knowl. Discov. 14(2):245-304.

[29] Fernández-Llatas C, Meneu T, Benedí JM, Traver V. [2010] Activity-based Process Mining for Clinical Pathways computer aided design. 2010 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC'10) 6178–6181.

[30] Weijters AJMM, Ribeiro JTS. [2011] Flexible heuristics miner (FHM). in IEEE Symposium on Computational Intelligence and Data Mining 310–317.