Evaluating the Effectiveness of Interactive Process Discovery in Healthcare: A Case Study

E. Benevento[1], P. M. Dixit[2], M. F. Sani[4], D. Aloini[3], and W.M.P. van der Aalst[4]

[1] University of Rome Tor Vergata, Rome, Italy
[2] Eindhoven University of Technology, Eindhoven, The Netherlands
[3] University of Pisa, Pisa, Italy
[4] Rheinisch-Westfälische Technische Hochschule (RWTH), Aachen, Germany
elisabetta.benevento@students.uniroma2.eu

Abstract. This work aims at investigating the effectiveness and suitability of Interactive Process Discovery, an innovative Process Mining technique, to model healthcare processes in a data-driven manner. Interactive Process Discovery allows the analyst to interactively discover the process model, exploiting his domain knowledge along with the event log. In so doing, a comparative evaluation against the traditional automated discovery techniques is carried out to assess the potential benefits that domain knowledge brings in improving both the quality and the understandability of the process model. The comparison is performed by using a real dataset from an Italian Hospital, in collaboration with the medical staff. Preliminary results show that Interactive Process Discovery allows to obtain an accurate and fully compliant with clinical guidelines process model with respect to the automated discovery techniques. Discovering an accurate and comprehensible process model is an important starting point for subsequent process analysis and improvement steps, especially in complex environments, such as healthcare.

Keywords: Interactive Process Discovery, Business Process Modeling, Healthcare, Process Mining.

1 Introduction and Background

Thanks to the pervasive adoption of Information Systems within healthcare organizations and the raising amount of patient and process-data, recent research has started focusing on data-driven approaches for investigating patient-flows through automatic or semi-automatic ways. Particularly, Process Mining (PM) has emerged as a suitable approach to analyze, discover, improve and manage real processes, by extracting knowledge from event logs [1]. Among the different PM perspectives, Process Discovery (PD) focuses on automatically discovering process models based on the event log, without using any apriori knowledge [1,2]. Of course, to gain significant outcomes, the event log should contain all the necessary information. A considerable number of PD techniques has been proposed by researchers for automatically discovering process models [1]. The most promising techniques for healthcare processes are

Heuristic Miner [3], Fuzzy Miner [4], Split Miner [5], and Inductive Miner [6], as they can handle noisy and incomplete event log [7,8]. Most of them produce formal models (Petri nets, transition systems, process trees, etc.), having clear semantics. In addition, there are available also several commercial tools (Disco, Celonis, QPR, ProcessGold, etc.) to support PD. They return process models that either have no formal semantics or correspond to so-called Directly-Follows Graphs (DFGs) that cannot express concurrency. These models provide valuable insights but cannot be used to capture the casual relationships of the activities in the process and draw reliable considerations [9].

PD is particularly critical for healthcare processes due to their intrinsic complexity, high variability and continuous evolution over time. [2,10]. Specifically, case heterogeneity typically leads to extract extremely complex, and often incomprehensible, process models, i.e. the so-called "spaghetti-like models" [1,11]. Besides, healthcare processes are highly dependent on clinicians' experience and expertise, i.e., they are knowledge-intensive [12], involving semi-structured and unstructured decision making. Such deep knowledge is not recorded in the event log, and, thus, it results difficult to elicit [13]. As a result, the mined models do not provide a meaningful representation of the reality, leading to a significant interpretation challenge for healthcare manager. To improve model quality, domain experts and analysts heuristically perform a refinement at the end of the discovery phase. Such a refinement is based on their knowledge, and it has turned out to be a time-consuming as well as iterative task [14].

Recently, new interactive PD approaches have been emerging, that allow to incorporate domain knowledge into the discovery of process models [13,15,16]. Combining domain knowledge and process-data may improve process modeling and lead to better results [17,18]. Interactive approaches are particularly useful in healthcare context, where physicians typically have a deep domain knowledge, whose integration within the process discovery phase can provide critical advances with respect to traditional automated discovery techniques [13,14].

This work aims at demonstrating the effectiveness and suitability of *Interactive Process Discovery* (IPD), an innovative interactive technique developed by Dixit [19], to model healthcare processes. IPD allows the user (i.e., the analyst or the expert) to interactively discover the process model, exploiting the domain knowledge along with the event log (Figure 1). In so doing, a comparative evaluation against the existing state-of-the-art process discovery techniques is carried out, in order to assess the potential benefits that domain knowledge brings in improving the quality and understandability of process models. The comparison is performed by using a real dataset from an Italian Hospital, in collaboration with medical staff.

The results confirm that IPD can outperform the existing process discovery techniques, providing a more accurate, comprehensible, and guideline compliant process model. Appropriate modeling of patient-flows may support healthcare managers in taking decisions related to capacity planning, resource allocation, and for making necessary changes in the process of care.

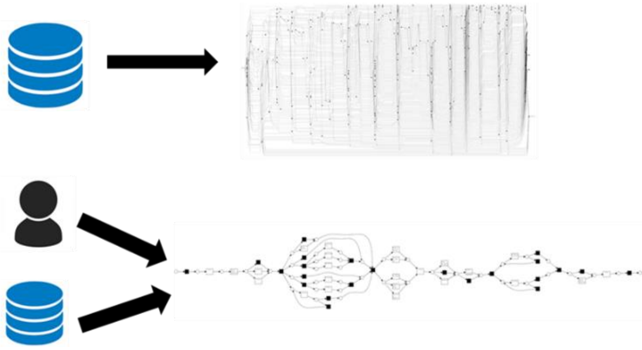

Fig. 1. Traditional automated Process Discovery (top) and Interactive Process Discovery (bottom) (process models are not intended to be readable).

2 Interactive Process Discovery

Interactive Process Discovery (IPD), developed by Dixit [19], is an innovative interactive technique for modeling knowledge-intensive processes based on the domain knowledge along with the event log. In IPD, the user (i.e., the analyst/expert) has total control over the discovery approach, and can model the process incrementally, at the desired complexity level, exploiting his deep knowledge. Information extracted from the event log is used to guide the user in making decisions about where to place a specific activity within the model. To enable the interactive discovery of a process model, the IPD miner uses the synthesis rules [19,20], which allow expanding a minimal synthesized net[1] by adding one transition and/or one place at a time. A brief description of IPD approach is reported here following (for more details see [19]). During the modeling phase, the user interacts with the synthesized net by applying arbitrarily three synthesis rules: (a) the abstraction rule, (b) the place rule and (c) the transition rule [20]. All possible applications of these rules are projected on the synthesized net, based both on the user interaction and on the information from the activity log[2].

More in detail, the user selects the activity to be added into the net from the activity log. Depending on the selected one, the status of the current synthesized net is updated. Specifically, IPD indicates to the user if the selected activity occurs before or after the other activities within the synthesized net. Alternatively, it highlights that the selected activity and the others in the network never take place at the same time. In so doing, IPD suggests where to place the activity, depending on the insights gained

1 A synthesized net is a free-choice workflow net containing a source place, a sink place, a start transition, and an end transition. For more details see [19,20].

2 An activity log is a multi-set (or bag) of sequences of activities. Every sequence of activities in the activity log is called an activity trace [19].

from the activity log. The user can decide to take assistance from the data or ignore the suggestion. The projected information can be based either on the eventually follows (precedes) relation or on the directly follows (precedes) relation, as desired by the user. The user labels the newly-added transitions in the synthesized net with an activity from the activity log. If the transition does not represent an activity, it is depicted as a silent transition. The activity label of the new transition is pre-selected by the user, after which the rule is applied.

3 Case Study: Objective and Methodology

In this work, IPD was applied to a real case of an Italian Hospital to show both the effectiveness of the approach and its suitability in a complex and knowledge-intensive environment. More in detail, we carried out a comparative evaluation against automated discovery techniques, to assess the potential benefit that domain knowledge brings in improving the quality of process models. The evaluation was performed in terms of accuracy and compliance with clinical guidelines.

The approach followed for the evaluation goes through three main steps: (a) data collection and preparation, (b) model building, and (c) model comparison (as depicted in Figure 2).

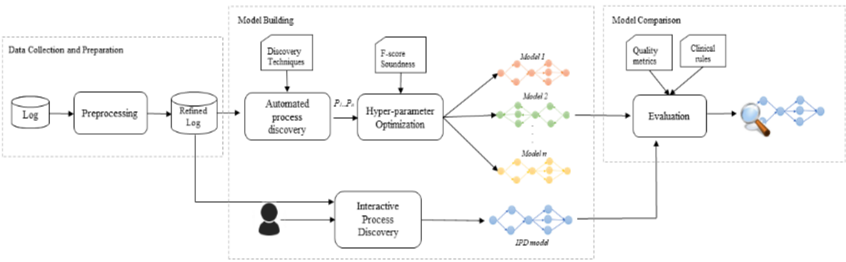

Fig. 2. Comparative evaluation approach.

3.1 Data Collection and Preparation

We collected and pre-processed data of all lung cancer patients treated by the hospital during the years 2014 and 2015. The management of lung cancer is complex and requires the integration of decisions made by practitioners from different disciplines. Decisions are mainly based on the practitioner's deep knowledge and expertise.

Data were mostly gathered from the Hospital Information Systems. The initial database consisted of 995 cases, 820 types of activities and more than 90,000 events. Before modeling, we decided to refine the raw event log, in order to guarantee its quality. As a matter of fact, it is directly related to the quality and the applicability of results. Data cleaning and preparation included: (a) outliers and incomplete cases removal, (b) low level activities aggregation, (c) less significant activities abstraction, (d) activity redundancy detection. As an example, we kept only the 21 most frequent activities to simplify the event log, since it contained a huge amount of different and fine-grained activities. In so doing, we aimed at building models with a comparable yet meaningful number of activities.

In the end, the refined event log consisted of 990 patient cases, 21 activities, and more than 14,000 events.

3.2 Model Building

Firstly, we applied IPD, as implemented in ProM 6.8, to extract the process model for lung cancer patients, with the collaboration of medical staff. To obtain the resulting model, on several occasions, we took assistance from insights of the event log gained via IPD (e.g., for positioning "radiotherapy" and "nuclear medicine" within the model). On some other occasions, we chose to ignore the information from the data, deeming it inadequate (e.g., for placing the "x-ray" within the model).

Following, among the state-of-the-art automated discovery techniques, we chose and applied the Inductive Miner (IM) [6], as implemented in ProM 6.8, and the Split Miner (SM) [5], as implemented in Apromore. As the 30 commercials tools produce Directly-Follows Graphs (DFGs) [9], we also applied the Directly-Follows Graphs Miner, as implemented in PM4Py [21].

As each discovery technique came with several parameters to be tuned, we optimized it by testing different parameter values, to find the best results. The optimization was carried by using a Rapid Miner extension, called RapidProM, and was based on the F-score metric [22] to find the solution with the optimal balance between fitness and precision. The F-score was computed on Petri nets since the measuring tools work only on Petri nets. Conversions of the process model in Petri nets were done using ProM's package.

3.3 Model Comparison

We measured and compared the quality of the model produced by IPD and of the best configurations of the automated discovery techniques in terms of accuracy and compliance with clinical guidelines.

To evaluate the accuracy of the process model, we experimented with two well-known metrics: fitness and precision [1], which both range between 0 and 1. The higher the fitness value, the more the model can replay the log. Conversely, the higher the precision value, the fewer behaviors (i.e., traces) are probable not to appear in the event log [1,9]. To compute fitness and precision, we resorted to the state-of-the-art alignment-based approaches described in [23,24]. Due to the trade-off between fitness and precision [22,25], we used the F-score as an evaluation metric, to take into account the ability of the model to equally fulfill and balance fitness and precision goals [21].

To assess the compliance of the model with the AIOM (Italian Association of Medical Oncology) guideline [26], we carried out a qualitative analysis with the collaboration of medical staff. Specifically, each process model was investigated from a "semantic point of view" and was evaluated on the capability to respect a set of medical rules. Specifically, the evaluation was based on the number of rules that were met by each process model. These medical rules were defined starting from the AIOM clinical guidelines and formalized by using a subset of (Declare) templates. Table 1 shows an overview and an interpretation of the templates that we considered. Each template provides a way to specify a dependency between two different classes of activities (e.g., a precedence constraint between the activities involved in the classes "surgery" and "medical examination") [27].

Table 1. Templates interpretation.

Type	Template interpretation based on the domain
Chain response (A,B)	**R1**: X-ray (B) should occur immediately after the Surgery (A)
Precedence (A,B) **&** **Not Succession (B,A)**	**R2**: Invasive diagnostic examination (B) should be preceded by radiological examinations (A) & (B) should not be followed by (A) **R3**: Surgery (B) should be preceded by invasive diagnostic examinations (A) & (B) should not be followed by (A) **R4**: Surgery (B) should be preceded by medical examinations (A) & (B) should not be followed by (A)
Precedence (A,B)	**R5**: If the removal of therapeutic aid (B) occurs, it should be preceded by the x-ray (A)
Init(A)	**R6**: The process should start with a general physical examination (A)
Existence (2,A)	**R7**: Lab test (A) should occur at least 2 times inside the process

4 Results

Table 2 reports the results related to the quantitative evaluation, i.e. the accuracy values obtained by IPD process model and by the best configurations for the automated discovery techniques. More in detail, the table summarizes, for each discovered model: (a) fitness, precision and F-score values; (b) the best configuration parameters provided as input (only for the automated discovery techniques).

Table 2. Quantitative evaluation of IPD process model and the best configurations for three representative automated discovery techniques.

PM technique		Best configuration parameters	Accuracy		
			Fitness	Precision	F-score
Interactive Process Discovery		-	0.70	0.64	0.67
Automated Discovery Techniques	Inductive Miner	1.0	0.59	0.71	0.64
	Split Miner	0.9 & 0.0	0.81	0.61	0.69
	DFG Miner	0.2	1	0.21	0.36

As shown in table 2, IPD, IM, and SM miners provide similar results in terms of F-score, unlike the DFG miner that is unable to balance fitness and precision values.

Note that we used the DFG model as a proxy for the models generated by commercial tools like Celonis, Disco, etc. As regards fitness and precision, all the techniques achieve different performance. Specifically, the DFG Miner strikes the best fitness with a value of 1, followed by the SM. However, the DFG Miner is less precise than the others, allowing behaviors not recorded in the event log. On the other hand, IM obtains a model that is slightly less able to reproduce the different behaviors in the log but more precise (with a value of 0.71). With IPD, experts could obtain a model with a quite high value of fitness, without penalization in precision. This is, definitely, a promising result in a knowledge-intensive domain such as the medical one.

Table 3 reports the scores obtained by the discovered process models in terms of satisfied rules.

Table 3. Number of rules satisfied by the models generated by the IPD, IM, SM, and DFG miners.

Rules	Interactive Process Discovery	Inductive Miner	Split Miner	DFG Miner
R1	1	0	1	0
R2	1	0	0	0
R3	1	0	0	0
R4	1	0	0	0
R5	1	1	1	0
R6	1	0	0	0
R7	1	0	1	1
Total value	7/7	1/7	3/7	1/7

Despite similar performance in terms of model accuracy, less than half of the rules were respected by the models generated by the IM, SM, and DFG miners, unlike IPD model. This is due to the fact that IM, SM, and DFG miners do not take the organizational information and the domain knowledge on the treatment process into account; as a result, their models fail to properly keep the structure of the process in line with clinical guidelines.

To better clarify this statement, let us drill down the behavior of each model with respect to rules R2 and R6. Figures 3,4,5 & 6 show the process models produced by IM, DFG miner, IPD, and SM respectively. In a healthcare context, some activities must follow a specific order of execution (see R1-R5 in table 3). For example, clinical guidelines suggest that invasive diagnostic procedures (e.g., bronchoscopy) must be executed immediately after radiological exams (x-ray and CT scan) and not vice-versa (R2), to confirm the diagnosis and evaluate the extent of the disease. Yet, IM, SM, and DFG miner models seem not to be able to capture this restrictive relationship, allowing also the inverse behavior for some process instances. Indeed, they use parallelism or exclusive choice with loops to represent the activities within the model. In such cases, the activities can take place in a different order from case to case, not respecting the restrictive condition (Figures 3, 4 & 6). On the other hand, in the process model produced by IPD, invasive diagnostic procedures are directly preceded by radiological exams (Figure 5). Furthermore, the lung cancer process should start with a general physical examination executed by the specialist (R6). Unlike IPD, the IM, SM, and DFG miner seem to violate this rule. Specifically, in both models produced by IM and DFG miner, the process may start with an activity other than the general physical examination (Figures 3 & 4). Similarly, in the SM model, the starting activity is the lab test, rather than the general physical examination (Figure 6). On the contrary, the correct order is properly captured in the IPD model (Figure 5).

These preliminary results confirm that leveraging domain knowledge and information recorded in the event log helps obtaining a process model that fully complies with the guidelines and is comparable in accuracy to the models produced by the automated techniques.

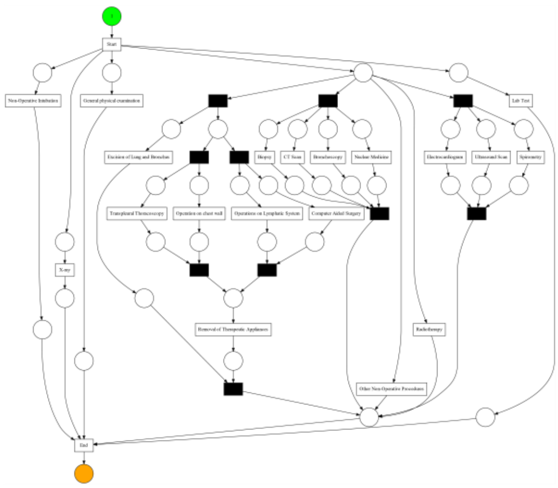

Fig. 3. Petri net for lung cancer patients generated by the Inductive Miner.

Fig. 4. Petri net for lung cancer patients converted from the DFG Miner (Note that the DFG Miner can be seen as a representative example of the discovery technique used by commercial systems).

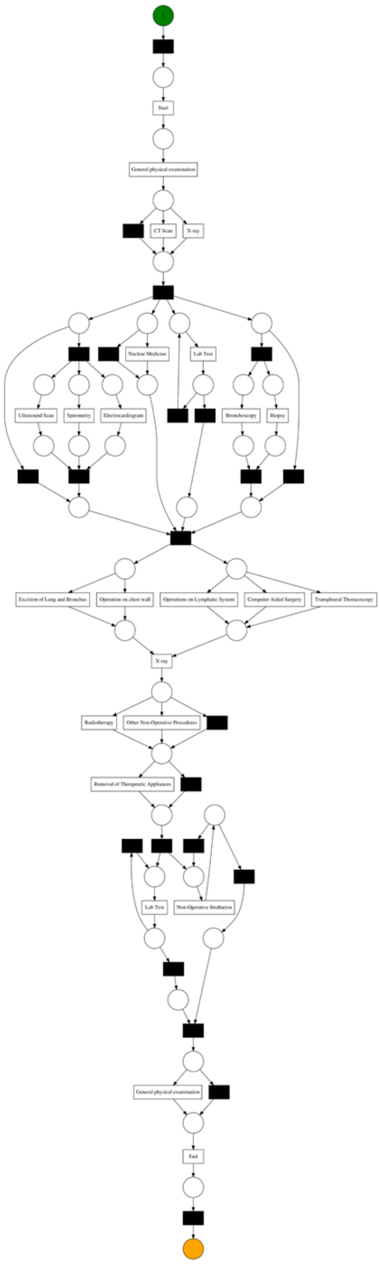

Fig. 5. Petri net for lung cancer patients generated by Interactive Process Discovery.

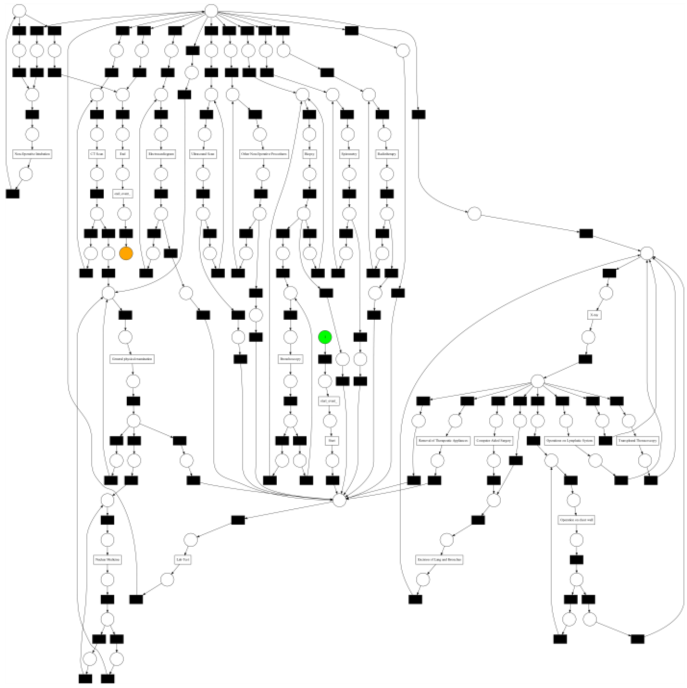

Fig. 6. Petri net for lung cancer patients converted from the Split Miner.

5 Discussion and Conclusions

In this study, we demonstrated the effectiveness and suitability of IPD [19] to model healthcare processes. IPD provides the analyst with a flexible way to interact with model construction, directly exploiting the domain knowledge along with the event log. Prior knowledge from domain experts represents a valuable resource in the discovery of process models, providing critical advances with respects to automated discovery techniques [13,19]. This is especially true in healthcare, where physicians typically have deep domain knowledge, not recorded in the event log, and, thus, difficult to elicit [14]. Therefore, both the automated discovery techniques fail in producing meaningful and comprehensible process models, resulting in a significant interpretation challenge for healthcare manager [19]. In contrast, IPD technique tries to structure the process data by using domain knowledge.

Our evaluation demonstrates that IPD can be used to obtain a guideline compliant process model, without penalizing its accuracy. Specifically, IPD achieves satisfactory results in terms of model accuracy, comparable to those of IM and SM. It also outperforms the DFG Miner. In addition, since IM, SM, and DFG Miner do not take the domain knowledge on the treatment process into account, their models fail to properly keep the structure of the process in line with clinical guidelines. On the contrary, IPD provides the ability to discover patient pathways that cover the most frequent medical behaviors which are regularly encountered in clinical practice, according to medical staff.

From a managerial viewpoint, discovering an accurate and comprehensible process model is an important starting point for subsequent process analysis and improvement steps, especially in complex environments, such as healthcare. Specifically, appropriate modeling of patient-flows may help healthcare managers to identify process-related issues (e.g., bottlenecks, process deviations, etc.) and main decisions.

While the initial experimental evaluation has provided satisfactorily results, in the future we aim at conducting a more extensive evaluation, replicating the study in different healthcare contexts to test the applicability and generalizability of IPD.

References

1. van der Aalst, W. M. P. Process Mining: Data Science in Action. Springer (2016).
2. Hompes, B., Dixit, P. M., & Buijs, J. Using Process Analytics to Improve Healthcare Processes. In Data Science for Healthcare (pp. 305-325). Springer, Cham (2019).
3. Weijters, A. J. M. M., & Ribeiro, J. T. S. Flexible Heuristics Miner (FHM). In IEEE SSCI 201: Symposium Series on Computational Intelligence-CIDM 2011: 2011 IEEE Symposium on Computational Intelligence and Data Mining (pp. 310–317) (2011).
4. Günther, C. W., & van der Aalst, W. M. P. Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 4714, 328–343 (2007).
5. Augusto, A., Conforti, R., Dumas, M., La Rosa, M., & Polyvyanyy, A. Split miner: automated discovery of accurate and simple business process models from event logs. Knowledge and Information Systems, 1-34 (2018).
6. Leemans, S. J., Fahland, D., & van der Aalst, W. M. P. Discovering block-structured process models from event logs containing infrequent behaviour. In International conference on business process management (pp. 66-78). Springer, Cham (2013).
7. Mans, R. S., Schonenberg, M. H., Song, M., van der Aalst, W. M. P., & Bakker, P. J. Application of process mining in healthcare–a case study in a dutch hospital. In International joint conference on biomedical engineering systems and technologies (pp. 425-438). Springer, Berlin, Heidelberg (2008).
8. Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. Process mining in healthcare: A literature review. Journal of biomedical informatics, 61, 224-236 (2016).
9. van der Aalst, W. M. P., De Masellis, R., Di Francescomarino, C., & Ghidini, C. Learning hybrid process models from events. In International Conference on Business Process Management (pp. 59-76). Springer, Cham (2017).
10. Rebuge, Á., & Ferreira, D. R. Business process analysis in healthcare environments: A methodology based on process mining. Information Systems, 37(2),99-116 (2012).
11. Diamantini, C., Genga, L., & Potena, D. Behavioral process mining for unstructured processes. Journal of Intelligent Information Systems, 47(1), 5–32 (2016).
12. Di Ciccio, C., Marrella, A., & Russo, A. Knowledge-Intensive Processes: Characteristics, Requirements and Analysis of Contemporary Approaches. Journal on Data Semantics, 4(1), 29–57 (2015).

13. Bottrighi, A., Canensi, L., Leonardi, G., Montani, S., & Terenziani, P. Interactive mining and retrieval from process traces. Expert Systems with Applications (2018).
14. Canensi, L., Leonardi, G., Montani, S., & Terenziani, P. Multi-level interactive medical process mining. In Conference on Artificial Intelligence in Medicine in Europe (pp. 256-260). Springer, Cham (2017).
15. Rembert, A. J., Omokpo, A., Mazzoleni, P., & Goodwin, R. T. Process discovery using prior knowledge. In International Conference on Service-Oriented Computing (pp. 328-342). Springer, Berlin, Heidelberg (2013).
16. Xu, X., Jin, T., Wei, Z., & Wang, J. Incorporating domain knowledge into clinical goal discovering for clinical pathway mining. In Biomedical & Health Informatics (BHI), 2017 IEEE EMBS International Conference on (pp. 261-264). (2017).
17. Dixit, P. M., Buijs, J. C., van der Aalst, W. M. P., Hompes, B. F. A., & Buurman, J. Using domain knowledge to enhance process mining results. In International Symposium on Data-Driven Process Discovery and Analysis (pp. 76-104). Springer, Cham (2015).
18. Mannhardt, F., de Leoni, M., Reijers, H. A., van der Aalst, W. M. P., & Toussaint, P. J. Guided Process Discovery–A pattern-based approach. Information Systems, 76, 1-18 (2018).
19. Dixit, P. M., Verbeek, H. M. W., Buijs, J. C. A. M., & van der Aalst, W. M. P. Interactive data-driven process model construction. In International Conference on Conceptual Modeling (pp. 251-265). Springer, Cham (2018).
20. Desel, J., & Esparza, J. Free choice Petri nets (Vol. 40). Cambridge university press (2005).
21. Berti, A., van Zelst, S. J., & van der Aalst, W. M. P. (2019). Process Mining for Python (PM4Py): Bridging the Gap Between Process-and Data Science. arXiv preprint arXiv:1905.06169.
22. De Weerdt, J., De Backer, M., Vanthienen, J., & Baesens, B. A robust F-measure for evaluating discovered process models. In IEEE Symposium on Computational Intelligence and Data Mining (CIDM-2011) (pp. 148-155) (2011).
23. Adriansyah, A., Munoz-Gama, J., Carmona, J., van Dongen, B. F., & van der Aalst, W. M. P. Measuring precision of modeled behavior. Information Systems and e-business Management, 13(1), 37-67 (2015).
24. Adriansyah, A., van Dongen, B. F., & van der Aalst, W. M. P. Conformance checking using cost-based fitness analysis. In 2011 IEEE 15th International Enterprise Distributed Object Computing Conference (pp. 55-64). IEEE (2011).
25. Sani, M. F., van Zelst, S. J., & van der Aalst, W. M. P. Improving process discovery results by filtering outliers using conditional behavioural probabilities. In International Conference on Business Process Management (pp. 216-229). Springer, Cham (2017).
26. Linee guida AIOM 2018. https://www.aiom.it/wp-content/uploads/2018/11/2018_LG_AIOM_Polmone.pdf. Last accessed 2019/04/15.
27. Maggi, F.M., Bose, R.J.C., & van der Aalst, W.M. P. A knowledge-based integrated approach for discovering and repairing declare maps. In International Conference on Advanced Information Systems Engineering (pp.433-448). Springer, Berlin, Heidelberg (2013).