Model discovery, and replay fitness validation using inductive mining techniques in medical training of CVC surgery

Gopi Battineni and Nalini Chintalapudi

E-Health and Telemedicine Centre,
School of Pharmaceutical Sciences and Health Products, University of Camerino,
Camerino, Italy, and

Francesco Amenta

E-Health and Telemedicine Centre,
School of Pharmaceutical Sciences and Health Products, University of Camerino,
Camerino, Italy and
Department of Studies and Research,
International Medical Radio Center Foundation (C.I.R.M.), Rome, Italy

Abstract

Medical training is a foundation on which better health care quality has been built. Freshly graduated doctors have required a good knowledge of practical competencies, which demands the importance of medical training activities. As of this, we propose a methodology to discover a process model for identifying the sequence of medical training activities that had implemented in the installation of a Central Venous Catheter (CVC) with the ultrasound technique. A dataset with twenty medical video recordings were composed with events in the CVC installation. To develop the process model, the adoption of process mining techniques of infrequent Inductive

Publishers note: The publisher wishes to inform readers that the article "Model discovery, and replay fitness validation using inductive mining techniques in medical training of CVC surgery" was originally published by the previous publisher of Applied Computing and Informatics and the pagination of this article has been subsequently changed. There has been no change to the content of the article.This change was necessary for the journal to transition from the previous publisher to the new one. The publisher sincerely apologises for any inconvenience caused. To access and cite this article, please use Battineni, G., Chintalapudi, N., Amenta, F. (2020), "Model discovery, and replay fitness validation using inductive mining techniques in medical training of CVC surgery" Applied Computing and Informatics. Vol. 18 No. 3/4, pp. 245-255. https://10.1016/j.aci.2020.01.001. The original publication date for this paper was 18/01/2020.

Miner (iIM) with a noise threshold value of 0.3 had done. A combination of parallel and sequential events of the process model was developed. Besides, process conformance was validated with replay fitness value about 61.1%, and it provided evidence that four activities were not correctly fit in the process model. The present study can assist upcoming doctors involved in CVCs surgery by providing continuous training and feedback on better patient care.

Keywords Process mining, iIM, Model discovery, Replay fitness, CVCs

Paper type Original Article

246

1. Introduction

Process mining is the latest research discipline, which support the business processes analysis based on event log data. The process mining research area has been used in the field of healthcare processes for the detection of process models, especially for conformance checking and analysis of social networks. It is highly focused on knowledge extraction from the data that was generated and stored in the IT systems [1]. Previously, the number of methods have been used to analyze the hospital processes, and Evidence-Based Medicine [2,3].

Recently, hospitals are using software systems to record patient data. These systems are indeed interacting with the real world. By default, these systems did not communicate with the outside world. As of this, the process model can be used to organize and explain how software systems can interacting with the real world [4]. It is an ideal method for learning these techniques by splitting event data recorded in hospital frameworks [5].

Mainly, in the medical training procedures, process mining has been utilized in various contextual investigations with better results [6,7]. Moreover, these algorithms can help doctors in the treatment process, thereby generating immediate medication [8–10]. Despite this, process model discovery and performance calculations can provide better chances in advantage gain from the information stored in hospital systems [8]. Using process-mining techniques in healthcare processes not only ensures process understanding, but also can improve the service quality, and positive patient feedback [9].

From cardiologists to health specialists, radiologists to anesthesiologists, and nurses to radiologic experts, there is a high demand for training sessions for medical students. There has been much exchange in the medical literature about the significance of compassion and doctor correspondence style in therapeutic practice. Some proofs explain a decrease in health care during the long clinical periods of medical school and continuing in an entire traineeship. Especially, surgeons are exposed to this problem due to the lack of idea in working place, and nonattendance of patient sessions. Ultimately, this is creating new issues for real doctors [11].

In this study, we aimed to present the importance of process mining techniques in the assessment of sequential activities involved in surgical training of Central Venous Catheter (CVC). The process mining tool (ProM) [12] with infrequent Inductive Miner (iIM) was adopted to process validation. Besides, this study addresses the conformance calculations. Therefore, it can be an easy way of finding the reasons for data deviation from the model. As per our best knowledge, this is the first study that uses conventional process mining techniques in medical training practices involved with CVC installation.

2. Methodology

2.1 Dataset

The medical training dataset from the conformance checking challenge (CCC.2019) [13] was adopted for this study. The particular dataset is composed of the medical training process, especially for medical students of the Pontifical Catholic University of Chile, to learn CVC installation with ultrasound [14]. This process suggests induction of catheter (tube) in a central vein, helping on conveying liquid or prescriptions to the patient, among different uses.

2.2 Observing event logs from data

The starting point of any process is data. Primarily, data selection was made in the form of XES (extensible event stream) conditions [15]. As discussed, data is measured and collected by medical students. For the adopted dataset, data labeled with 20 instances of 29 individual activities. Each instance is involved with a separate ten video recordings (each recording was conducted in two times) of CVC installation. Each record has checked and verified by the corresponding tutor in the university.

2.3 Mapping of CSV to XES

Process mining involved with event logs and each event corresponds to a single activity at a particular instance. This event data either available from the hospital database (for example, patient data) or spreadsheets or comma-separated value (CSV) files. However, because of many process mining tools adopt XES standard format, it was demanded to convert CSV to XES format (Figure 1).

2.4 Infrequent inductive miner (iIM)

Several tools are available for process mining and algorithms for model generation, tables, and data analysis. However, ProM is an open-source tool that consists of many methods for process model development. ProM version 6.9, coupled with iIM was employed to ensure the model soundness [16]. Inductive miner is a discovery algorithm, depending on the outcomes from soundness in-process models (i.e., Petri-nets). Petri nets are useful for model visualization in the process mining [17]. In the end, model validation was done with a threshold noise level of the model was equal to 0.3.

2.5 Replay log fitness to conformance check:

Replay log fitness defines the model characteristics and also whether a model can replay the observed behavior or not [16]. In simple words, process conformance explains the gap between the fitting model with traces (trace is a series of activities) and the actual model (which is in mind). These alignments are fundamental techniques to calculate the replay fitness and conformance checking and helped to find the data deviation in the model. Once the Petri-net model was generated, then the model coupled with a log file can estimate the replay fitness.
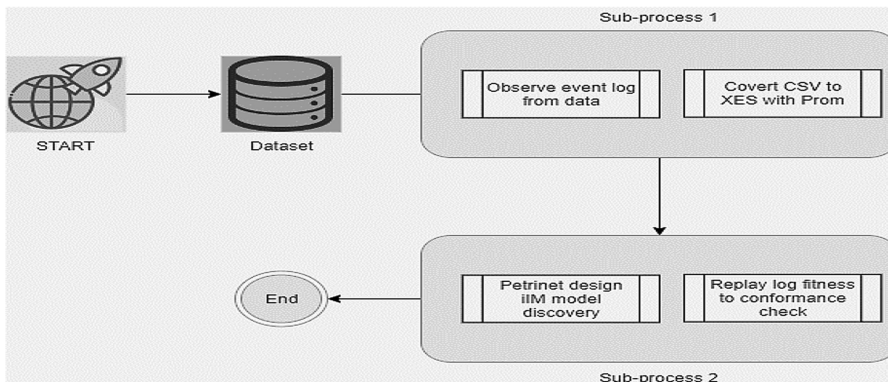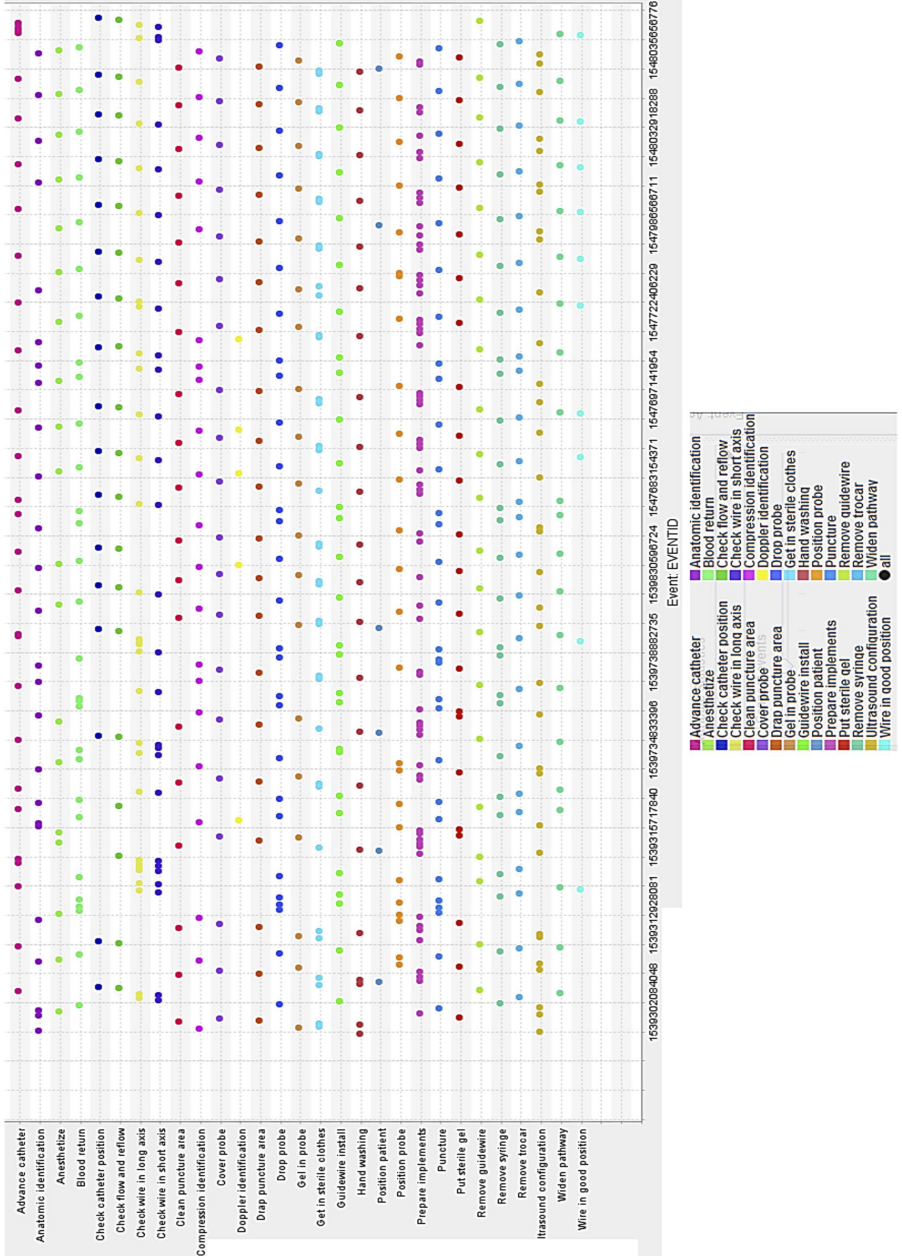

Figure 1.
The sequential methodology flow diagram.

Figure 2.
Sample event
distribution over all the
activities.

3. Results

3.1 Dataset outcomes

The follow-up to data mapping, event distribution with total activities in CVC installation was observed (Figure 2), and resource occurrence portions (high to low) were estimated (Table 1).

3.2 Mining of CVC schemas

In health care, process mining was involved in the control stream, execution, conformance, and reliability [18]. As of this, after the log extraction, different Petri-net models were generated by adjusting distinct noice values. Ultimately, the model linked with primary activities was generated with a threshold noise level equal to 0.3 (Figure 3).

In the first section of the process, a parallel flow of event classes can be observed. The method initiates with three parallel events called get in sterile cloths, anesthetize, and doppler identification. It defines that there is no strict rule in the activity initiation. Once the doctor in sterile cloths, the next steps would follow hand washing and keeping the patient in the right position. After that, the process follows two flows: one starts with a clean puncture area followed by a drop puncture area, and implant preparation indicating ending the process. Other flows start with a puncture, followed by blood return and dropping probe. After these steps, rest of the flow followed with sequential events of syringe removal, guide ware install, trocar removal, wire checking with the short and long axis.

3.3 Conformance checking

The preliminary model outcome of conformance check and synchronous activity flow represented with green and asynchronous activity flow others with red-colored edges can be observed (Figure 4). Trace fitness value was validated with 61.1%, and 38.9% of model events were not accurately fit in the given trace; also, perform activities in an asynchronous manner. Total four activities were not performing according to the clinical practice plan. Consequently, this was affecting the overall model performance. For example, Activities of anesthesia, trocar removal, and short-axis wire check were conducted (17 times) in a synchronous manner, and two times the model move had observed (Table 2). Therefore, 2 out of 19 instances, mentioned activities were not found in the traces. Similarly, long axis wire check activity was not observed a single time in the observed traces.

4. Discussion

The present study is aimed to discover process model design, and data deviation causes between activities in medical training of CVC installation. In this study, we adopted the iIM

N	Class	Occurrences (absolute)	Occurrences (relative)
1	R_21_1F	180	12,912%
2	R_31_1G	154	11,047%
3	R_32_1H	154	11,047%
4	R_48_2D	150	10,76%
5	R_45_2A	138	9,9%
6	R_33_1L	134	9,613%
7	R_14_1D	130	9,326%
8	R_13_1C	126	9,039%
9	R_46_2B	122	8,752%
10	R_47_2C	106	7,604%

Table 1.
Resources (records corresponding with occurrence and percentages).

Figure 3.
Process Model with
medical training
activities involved in
CVCs installation
(i.e. White circles and
black boxes symbolize
places and transitions
respectively).

Figure 3.

Figure 4.
Conformance check outcome model (Green color-coded activities represent correct synchronous with traces, and red-bordered events represent asynchronous trace alignment). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).
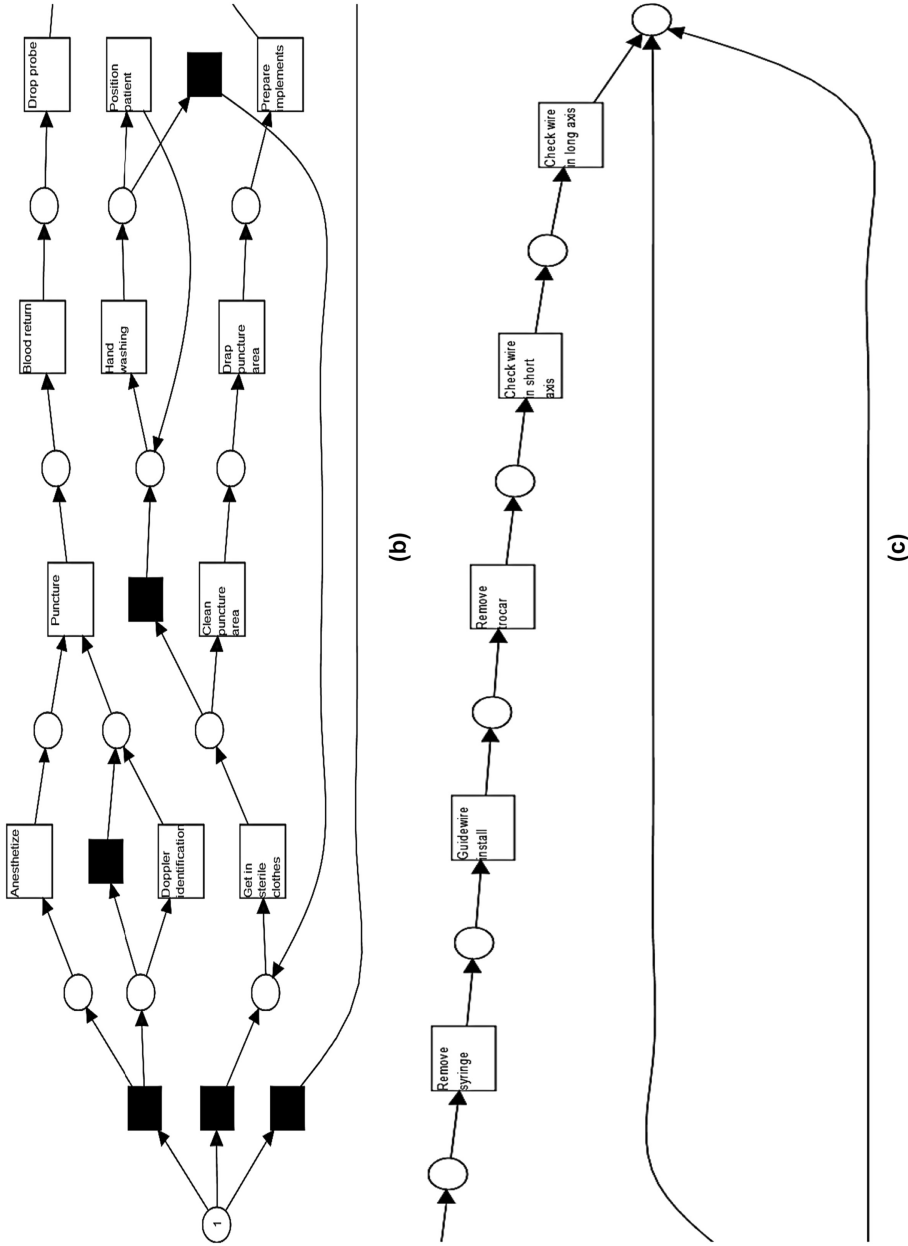
technique to develop a Petri net model to highlight the medical activity sequences. Besides, a conformance check involved in calculating the model fitness was well explained. The synchronous behavior of 61.1% event logs, and asynchronous behavior of 38.9% event logs was observed. Four primary activities affecting the model performance (Table 2) are observed. Results revealed that it is essential to have immediate steps for log fitness improvement. A study of generating enhanced structural medical process [19], helped to gain a better understanding of various hospital models. In real clinical practices, workforce setting to manage the CVC position under elective and planned conditions. Due to the involvement of advanced computing methods in healthcare, the requirement of central venous access should be fast, lifesaving measures, and the specialist should do the CVC position [20,21]. Our findings can contribute to understanding the insight process behavior of medical training for enabling doctors to finish the surgical process on time.

To the best of our knowledge, it is the first study on the conformance check challenge on CVC installation. In the study of [14], an explanation on the use of process mining strategies to find out feedback on different interests of the surgery process was done. The main limitation of this study was not to disclose the primary activity involvement and their contribution to the conformance check. However, the current research avoids this limitation by including primary activities of CVC surgery. We observed three parallel flows in the first stage of operation. The doctor might choose individual flow according to the activity plan. After that, process flow was followed by the synchronous flow of activities until the end of the surgery. This study is in line with [22] that united with the proposed methodology of process patterns detection in structured learning practices.

Despite this, another limitation is a low process instance rate, which might hamper the supposition of outcomes to the overall medical students' population. However, we believe that this is an initial step to conduct process model schemas in the particular program of medical training. Another important issue was low replay fitness value (61.1%), as per Ref. [23]; if any model possesses low replay fitness value, the model soundness (i.e., process model able to reach the end state without any error) will be highly affected. These issues could be addressed in future studies for validating the process mining usage in the real world.

5. Conclusions
In health care, designing high-level learning practice could be essential for the acquisition of better results in surgeries. Supplement of primary training for medical trainees could enable them to develop specific surgery methods in a short time. This paper mainly discussed the usage of process mining in the medical interest of CVC installation. An Inductive mining technique was employed in finding activity sequences at surgery time. The outcome model represents a structural view of the process model with a combination of sequential and parallel events. Besides, the present study limited to low model fitness values, and it could be addressed in future studies by involving advanced process mining techniques.

Activity	Synchronous moves ratio	Move model ratio	
Anesthetize	17/19	2/19	**Table 2.**
Remove Trocar	17/19	2/19	Activities that not fit
Check wire in short axis	17/19	2/19	correctly in the process
Check wire in the long axis	18/19	1/19	(model coupled with move log, and move model ratio).

References

[1] A.J.M.M. Weijters, W.M.P. van der Aalst, Process mining: a research agenda, Comput. Ind. (2004).

[2] W. Van Der Aalst, T. Weijters, L. Maruster, Workflow mining: Discovering process models from event logs, IEEE Trans. Knowl. Data Eng. (2004).

[3] K. Kirchner, N. Herzberg, A. Rogge-Solti, M. Weske, Embedding Conformance Checking in a Process Intelligence System in Hospital Environments, 2013.

[4] W.M.P.M.P. van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes. 2011.

[5] W.S. Yang, S.Y. Hwang, A process-mining framework for the detection of healthcare fraud and abuse, Expert Syst. Appl. (2006).

[6] Á. Rebuge, D.R. Ferreira, Business process analysis in healthcare environments: A methodology based on process mining, Inf. Syst. (2012).

[7] E. Rojas, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Process mining in healthcare: A literature review, J. Biomed Inform. (2016).

[8] I. Yoo, et al., Data mining in healthcare and biomedicine: A survey of the literature, J. Med. Syst. (2012).

[9] 'PROCESS MINING IN HEALTHCARE, A case study', in Proceedings of the First International Conference on Health Informatics 2008.

[10] I. Yoo, et al., Data mining in healthcare and biomedicine: A survey of the literature, J. Med. Syst. (2012).

[11] J.L. Han, T.N. Pappas, A review of empathy, its importance, and its teaching in surgical training, J. Surg. Educ. (2018).

[12] B.F. van Dongen, A.K.A. de Medeiros, H.M.W. Verbeek, A.J.M.M. Weijters, W.M.P. van der Aalst, 'The ProM Framework: A New Era in Process Mining Tool Support', 2005.

[13] 4TU – Dataset: Conformance Checking Challenge 2019 (CCC19)'. [Online]. Available: https://data.4tu.nl/repository/uuid:c923af09-ce93-44c3-ace0-c5508cf103ad. [Accessed: 24-Dec-2019].

[14] R. Lira, et al., Tailored Process Feedback Through Process Mining for Surgical Procedures in Medical Training: The Central Venous Catheter Case, 2019, pp. 163–174.

[15] W.M.P. Van Der Aalst, M.H. Schonenberg, M. Song, Time prediction based on process mining, Inf. Syst. (2011).

[16] R. Conforti, M. La Rosa, A.H.M. Ter Hofstede, Filtering out infrequent behavior from business process event logs, IEEE Trans. Knowl. Data Eng. (2017).

[17] T.M. Chen, J.C. Sanchez-Aarnoutse, J. Buford, Petri net modeling of cyber-physical attacks on smart grid, IEEE Trans. Smart Grid (2011).

[18] R.S. Mans, W.M.P. Van Der Aalst, N.C. Russell, P.J.M. Bakker, A.J. Moleman, Process-Aware Information System Development for the Healthcare Domain – Consistency, Reliability, and Effectiveness', in Lecture Notes in Business Information Processing, 2010.

[19] S. Montani, G. Leonardi, S. Quaglini, A. Cavallini, G. Micieli, Improving structural medical process comparison by exploiting domain knowledge and mined information, Artif. Intell. Med. (2014).

[20] A.J. Ullman, et al., Dressings and securement devices for central venous catheters (CVC), Cochrane Database Syst. Rev. (2013).

[21] V. Chopra, J.C. O'Horo, M.A.M. Rogers, D.G. Maki, N. Safdar, The risk of bloodstream infection associated with peripherally inserted central catheters compared with central venous catheters in adults: a systematic review and meta-analysis, Infect. Control Hosp. Epidemiol. (2013).

[22] M. Cameranesi, C. Diamantini, L. Genga, D. Potena, 'Students' careers analysis: a process mining approach, Assoc. Comput. Mach. (2017).

[23] J.C.A.M. Buijs, B.F. Van Dongen, W.M.P. Van Der Aalst, 'On the role of fitness, precision, generalization and simplicity in process discovery', in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012.

Corresponding author
Gopi Battineni can be contacted at: gopi.battineni@unicam.it