Tailored Process Feedback through Process Mining for Surgical Procedures in Medical Training: the Central Venous Catheter case

Ricardo Lira[1], Juan Salas-Morales[1], Rene de la Fuente[2], Ricardo Fuentes[2], Marcos Sepúlveda[1], Michael Arias[3], Valeria Herskovic[1], and Jorge Munoz-Gama[1]

[1] Computer Science Department, School of Engineering
[2] Department of Anesthesiology, School of Medicine
Pontificia Universidad Católica de Chile (Chile)
{rlira2,jisalas1,jmun}@uc.cl - {marcos,vherskov}@ing.puc.cl
{rdelafue,rfuente}@med.puc.cl
[3] Department of Business Computer Science
Universidad de Costa Rica (Costa Rica)
michael.arias_c@ucr.ac.cr

Abstract. In healthcare, developing high procedural skill levels through training is a key factor for obtaining good clinical results on surgical procedures. Providing feedback to each student tailored to how the student has performed the procedure each time, improves the effectiveness of the training. Current state-of-the-art feedback relies on Checklists and Global Rating Scales to indicate whether all process steps have been performed and the quality of each execution step. However, there is a process perspective not successfully captured by those instruments, e.g., steps performed but in an undesired order, part of the process repeated an unnecessary number of times, or excessive transition time between steps. In this work, we propose a novel use of process mining techniques to effectively identify desired and undesired process patterns regarding rework, order, and performance, in order to complement the tailored feedback of surgical procedures using a process perspective. The approach has been effectively applied to analyze a real Central Venous Catheter installation training case. In the future, it is necessary to measure the actual impact of feedback on learning.

Keywords: process mining, healthcare, feedback, medical training, surgical procedures

1 Introduction

The development of procedural skills is critical for physicians of any specialty. Better technical skills are associated with better clinical outcomes [8] and the absence of them is the most important factor in errors derived from the operator in healthcare [13]. Historically, procedural skills have been taught in daily clinical work, in a master-apprentice model, assuming sufficient exposure time to obtain them. However, there are a number of complexities that increasingly make this model more difficult: health system efficiency considerations, time constraints on clinical training activities, and patient safety [19]. In particular, [19] mentions some examples of efficiency issues: resident work hour restrictions have reduced the exposure of residents to their surgical mentors; and changes in reimbursement and insurance and legal issues have introduced productivity constraints on the surgical procedures. Therefore, training in simulation environments prior to contact with patients has extended as an effective practice to achieve positive effects on the learning process [5]. However, this teaching methodology has a high cost [20] and presents aspects that are not fully resolved, including how to give effective feedback to the students.

Feedback in clinical education of procedural skills is defined as the delivery of specific information on the comparison between the student's performance and a standard [15]; it ensures that certain standards are met, and promotes learning [4]. Feedback is effective when it is used to promote a positive and desirable development [2]. In the case of procedural skills taught in a simulation environment, feedback can be delivered by an instructor, partner or computer, either during or after the simulation activity. However, standard oral feedback has some drawbacks, since it depends on the availability of a person who is trained in the procedure, which is usually a difficult to obtain and expensive resource. In addition, it is an essentially subjective opinion of the evaluator.

To establish when an apprentice has reached an acceptable level of competence that guarantees patient safety, various tools have been developed. The two most commonly used are Checklists and Global Rating Scales (GRS) [11]. Checklists break down the procedure into a series of steps, and check whether the student has performed each step. Meanwhile, GRS consider the evaluation of the student's performance in different areas. In both cases, at the end of each training session, they allow to provide feedback to the student about which steps/areas deserve a greater attention. Checklists have the limitation of giving similar weights to different errors in the execution of a procedure, even though some of these have greater implications for the clinical outcome and patient safety [14]. On the other hand, GRS provide a more qualitative evaluation, but have the limitation that their reliability is dependent on the characteristics and training of the evaluators [3].

The performance of a surgical procedure can be seen as a process [16], i.e., a set of activities (procedure steps) and events that are executed in a specific order so as to achieve a certain goal. A process-oriented feedback seeks to emphasize the relevance of following this orderly sequence of activities, identifying deviations such as: rework, execution of activities in a different order than desired, slow execution of activities, or slow transition times between activities. Process Mining [1] is an emerging discipline that allows analyzing the execution of a process based on the knowledge extracted from event logs created from the data stored in information systems. The event logs record the execution of the different activities in which a process can be broken down.

In this article, we propose the novel use of process mining in order to complement the tailored feedback of surgical procedures using a process perspective. In particular, we believe it is possible to identify when the student repeats some activities (rework), when the student performs some activities in an incorrect order, or when it takes too long to perform an activity or a transition between two consecutive activities. To the best of our knowledge, this is the first application of process mining in medical procedures training.

The proposed method has been validated by applying it to a course, taught by the simulation center of the Pontificia Universidad Católica de Chile, where students learn the procedure for installing the Central Venous Catheter with ultrasonography. In this course, the traditional method of feedback delivery is immediate oral feedback by the instructor, along with an evaluation based on Global Rating Scales and Checklists.

2 Objectives and Context

This article has two main contributions. First, we propose a novel method for applying process mining techniques to effectively identify desired and undesired process patterns regarding rework, order, and performance, in order to complement the tailored feedback of surgical procedures using a process perspective. Second, we illustrate how this method was applied to a real Central Venous Catheter installation training case.

2.1 Objectives

Our main research objective (**O**) is to propose how *process mining* can be used to identify desired/undesired process patterns as part of the tailored feedback on medical procedural training. It can be broken down into specific objectives:

O1: To identify desired/undesired process patterns regarding *rework*.
O2: To identify desired/undesired process patterns regarding the *order* in which activities are performed.
O3: To identify desired/undesired process patterns regarding *performance*.

2.2 Central Venous Catheter installation training case

The simulation center at the School of Medicine of the Pontificia Universidad Católica de Chile, developed a training program for 42 first-year residents of anesthesiology, emergency medicine, cardiology, intensive medicine and nephrology, in the context of the research "Simulation-based training program with deliberate practice for ultrasound-guided jugular central venous catheter placement" [6]. This program is developed in three stages:

I. Online Instruction and PRE recording: three online classes are available through a web platform, each with mandatory and complementary readings. At the end of this stage, a written evaluation is taken and a recording of a first procedure execution is made (identified as PRE video).

II. Demonstration Session: a demonstration of the entire procedure of installation of a Central Venous Catheter (CVC) in a "Blue Phantom Torso"[4] is given by an expert to the entire group of residents. In addition, the 4 stations of deliberate practice are presented: a) preparation of the ultrasonography equipment, the patient and the work tools; b) handling ultrasonography equipment; c) venous puncture with ultrasound guidance; d) catheter installation and fixation.

III. Deliberate Practice: residents must complete four deliberate practice sessions accompanied by an instructor who supervises and delivers immediate feedback in the development of stations described in stage II.

After the end of the course, a second video (identified as POST video) of the procedure of installation of a CVC with ultrasonography is recorded for each resident, which is used to evaluate the training program.

Parallel to this training session, we recorded videos (identified as EXP videos) of the execution of the same procedure under the same conditions by different professionals from the anesthesiology division, with at least 5 years of clinical practice and experience in the installation of CVC with ultrasonography.

3 Method

Unlike more classical process mining methodologies, in our approach event logs are not generated automatically from the execution of the procedure. Instead, our method uses an observer-based approach [12], i.e., observation performed by a human observer. In this case, event logs were generated based on the off-line observation, by medical specialists, of the aforementioned recorded videos.

The proposed method is inspired by the process mining PM2 methodology [7], which has been previously used in the healthcare domain [17, 18] and it is suitable for the analysis of both structured and unstructured processes. The proposed method is decomposed into five stages (see Figure 1): (1) video recording - data are extracted from the videos recorded for both students and experts; (2) video tagging - the videos were tagged by two medical doctors, identifying for each case (each execution of the CVC installation), activities (procedure steps) and timestamps (time elapsed since the beginning of the procedure); (3) event log generation - tagging information is used to generate an event log containing the data of all executions; (4) model discovery - process mining discovery algorithms are applied to the event log in order to describe the observed behaviour of the procedure; and, (5) model analysis - the discovered process models are analyzed in order to generate feedback for each student.

3.1 Video recording stage

In this stage, we extracted trace data from different recorded videos that register how the procedure was executed by each of the students. The videos were organized into three categories: PRE: executions previous to the training program, POST: executions after the training program, and EXP: executions by experts.

[4] http://www.bluephantom.com

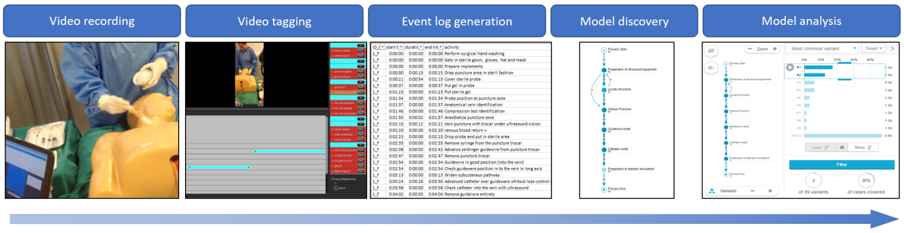

Fig. 1. Stages of the proposed method.

3.2 Video tagging stage

In this stage, each video is tagged by two medical doctors. VCode Vdata is used as tagging software [10]. A set of activities that are characteristic of the CVC procedure are used as possible labels for each of the activities observed in the videos. The result of the tagging process is a plain text file for each video; each of its rows displays information on: activity name, start time and duration time in hundredths of a second, and finally an optional field of notes.

3.3 Event log generation stage

In this stage, we created the event log that is used by process mining algorithms. The event log is a comma separated value file that included a row for each of the events tagged, grouping in a single file the data gathered from all the videos. It contains the following columns: case id (each video), event (activity name), start and end timestamps (both with a granularity at the second level), and an observation field. In addition, for each video, the following fields are recorded: performer (participant id), type of performer (student or expert), category (PRE, POST or EXP), and success or failure in the execution of the procedure.

3.4 Model discovery stage

We processed the event log with a discovery algorithm to obtain a process model representing the behaviour of each student when performing the procedure. In the PM literature, there is a wide range of discovery algorithms [1]. We selected the Celonis algorithm and its implementation in the Celonis commercial tool[5]. This algorithm is based on the Fuzzy algorithm concept [9] combined with some characteristics from the family of Heuristic algorithms [1], providing process models that are easy to interpret for an interdisciplinary audience. Moreover, Celonis tool also integrates a set of filtering options to explore the process data interactively and to address the specific objectives we have.

[5] Celonis tool: http://www.celonis.com/en/product

3.5 Model analysis stage

Celonis was used to identify aspects of the execution of the process that can be delivered as feedback to the student, to guide their learning. Specifically: 1) identify rework, i.e., when the student repeats one or more activities in the execution of the procedure; 2) identify the execution of activities in an incorrect order of execution, comparing it with the order in which the experts perform it; 3) analyze the student's performance, including duration of activities and transition time between them, comparing it with the performance of the experts.

We consider two key features to create custom views that can be used to provide process-oriented feedback to the students:

Filter: filters can be applied to the event log in order to obtain more specific process models. We use three kind of filters:

Activity selection filters: They allow to exclude some activities from a process model. For example, we distinguished two type of activities: action activities are those that are performed in order to install the CVC, and checking activities are those that are performed in order to verify whether some critical activities produce the desired outcome. In some models, we exclude the checking activities.

Case selection filters: They allow to create a process model using only some process instances. For example, we use these filters to create process models that display only the activities performed by the student we want to give feedback to, either in the PRE or POST scenario, and to create a process model that display only the activities performed by the group of experts (EXP).

Collapsing filters: They allow to group some activities in a process model, so that they are represented by a single element in the model. For example, some phases of the CVC installation are regularly performed well by most participants. Therefore, all activities corresponding to one of those phases can be clustered in a single cluster element.

Compose view: We compose views that include different models of a process based on the data loaded from an event log. Each of those models can be created by applying different filters to the event log in order to obtain more specific process models.For example, Figure 2 shows a composition of three views, displaying the execution of the procedure by a student before/after the training, compared to the execution of the procedure by an expert.

4 Results

The process-oriented feedback for a student who is learning about a surgical procedure is delivered through different process models that show patterns in which their performance is compared to the desired behavior. This information is useful for students because it helps them to focus their attention and effort, so as to avoid making mistakes in future executions. It can be a guide for future sessions of deliberate practice, which allows focusing efforts on simpler and independent tasks, e.g., puncturing the vein with ultrasonography or passing the guidewire, thus simplifying the training to some steps that are difficult due to rework, lack of order, or slowness in execution. Different views and filters were applied to address the three specific objectives:

***O1**. To identify desired/undesired process patterns regarding* rework*:* For this objective, we compose a view with three models (as shown in Figure 2). The first two models display how the student performed the procedure before and after the training (PRE and POST). The third model displays how an expert would perform the procedure. Since the purpose is to illustrate the occurrence of reworks in the action activities, checking activities are excluded. This view allows to provide feedback about which activities the student repeated (**O1.1**); to comment on the performance's evolution, by comparing the reworks observed in PRE versus POST (**O1.2**); finally, by including the execution of an expert, it is possible to compare the student's performance against the desired outcome (**O1.3**). In the example shown in Figure 2 a rework in the trocar installation can be observed.

***O2**. To identify desired/undesired process patterns regarding the* order *in which activities are performed:* To achieve this objective, we compose a view with three models (Figure 3). The first two models display how the student performed the procedure before and after the training (PRE and POST). The third model displays how an expert would perform the procedure. Since the purpose is to illustrate problems in the order in which activities are performed, those phases that are correctly performed by most participants are excluded. This view allows to provide feedback about which activities the student performed in the wrong order, or activities that were not performed at all (**O2.1**); to comment on the performance's evolution, by comparing the performance PRE versus POST (**O2.2**); finally, it is also possible to compare the student's performance against the desired order as performed by an expert (**O2.3**). In the example shown in Figure 3, *Remove trocar* and *Check guidewire* activities are performed in the opposite order to the desired one.

***O3**. To identify desired/undesired process patterns regarding* performance*:* To achieve this objective, we compose a view with three models (as shown in Figure 4). The first two models display how the student performed the procedure before and after the training (PRE and POST) including the time required to perform each activity and the time elapsed during the transition between two consecutive activities. The third model displays the average time it takes the group of experts (EXP) to perform the procedure. This view allows to provide feedback about which activities were performed too slow (**O3.1**), or when the student hesitated taking too much time between activities (**O3.2**). In this case, it is always useful to have as a reference the performance of the experts. In Figure 4, it can be observed in the PRE model that *Puncture trocar* was performed too slow and that the transition time between *Remove trocar* and *Widen pathway* took too long.

Fig. 2. Process model of student *16F69F* in PRE (left) and POST (center), and a generic expert EXP (right). Numbers show activity frequency per case. Activity names are shortened and only action activities are shown for readability reasons. Regarding **O1.1**, the PRE model shows rework during the phase of venous puncture with trocar: the student unsuccessfully performs a first trocar installation, then perform a successful one on the second iteration. Moreover, during the second iteration, the probe is not properly dropped (in the sterile zone) closing the door for a possible third iteration without sterilizing everything again. Notice that, the exact path followed in each iteration can be analyzed using the case animation feature of the tool. Regarding **O1.2**, POST model is able to capture that the same student does not perform any rework, and the process match exactly the reference process performed by the EXP (**O1.3**).

Fig. 3. Process model of student *12F67F* in PRE (left) and POST (center), and a generic expert EXP (right). Activities in *set trocar* and *check catheter* phases are clustered for the sake of readability (they show no difference in order between PRE, POST and EXP). Activity names are shortened and both action and checking activities are shown. Regarding **O2.1**, the PRE model shows the student checked the guidewire and then removed the trocar; however, it is desirable to do these activities in the opposite order, because removing the trocar may affect the position of the guidewire. Regarding **O2.2**, POST model is able to capture that the student learned to perform the activities in the right order, and the process match exactly the reference process for the procedure performed by the EXP (**O2.3**).

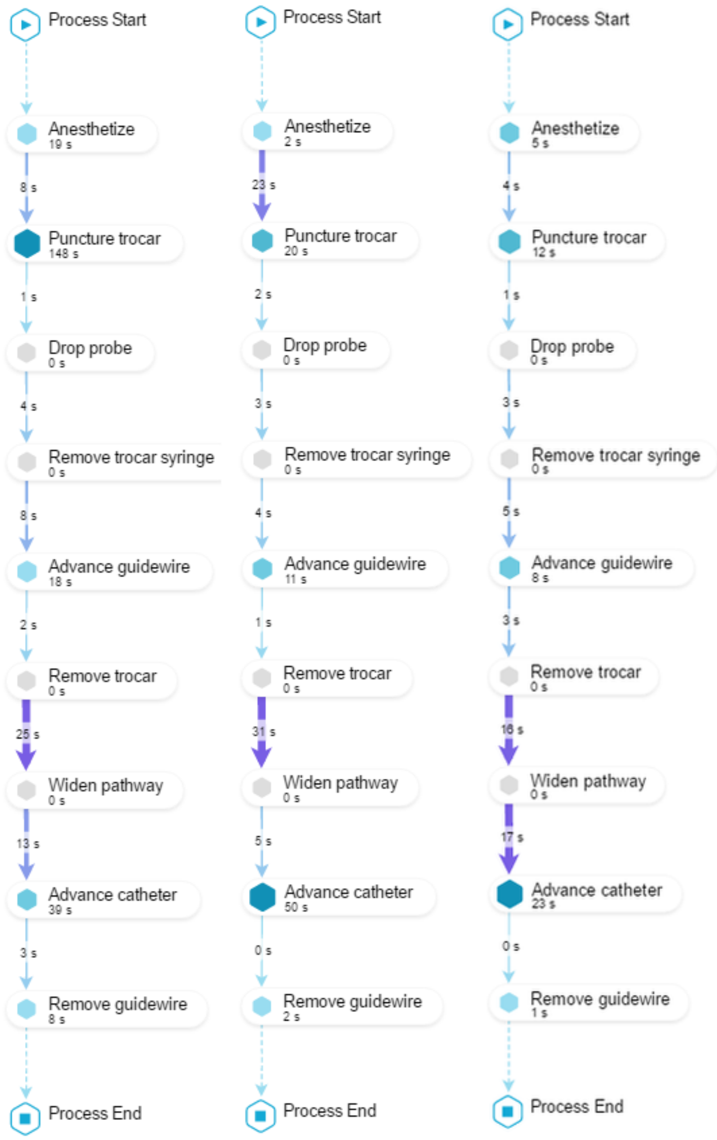

Fig. 4. Process model of student *45D59F* in PRE (left) and POST (center), and the average of the 8 experts in EXP (right). Activity names are shortened and both action and checking activities are shown. Regarding **O3.1**, it is possible to highlight *Puncture trocar* was performed too slow in the PRE model (notice the symbol of this activity is larger and darker). Regarding **O3.2**, it can be observed that the transition time between *Remove trocar* and *Widen pathway* is too long (notice the arrow between these activities is wider and darker). Regarding **O3.3**, it is possible to highlight the duration of *Puncture trocar* in the POST model got close to the EXP model. However, the time between *Remove trocar* and *Widen pathway* is still too long (**O3.4**).

It is also possible to provide feedback about student's evolution from PRE to POST regarding the duration of activities (**O3.3**) and the time between transitions (**O3.4**). It could happen that some activities (or times between activities) have similar execution times to the average time of the experts, but others are still not close to the average time of the experts. It is then possible to highlight positive aspects, and others where there is still room for improvement. In the example shown in Figure 4, it is possible to highlight in the POST model that the duration of *Puncture trocar* got close to the EXP model, but the transition time between *Remove trocar* and *Widen pathway* is still too long (**O3.4**).

5 Conclusions and Future Work

In this article, it has been shown that process mining can be used to provide a process-oriented feedback to students who are learning procedural skills, by effectively identifying desired and undesired process patterns regarding rework, order, and performance, in order to complement the tailored feedback of surgical procedures. The approach has been effectively applied to analyze a real CVC installation training case. To the best of our knowledge, this is the first application of process mining in medical procedures training. It opens a novel approach to the analysis of training programs by generating tailored feedback to students. This approach is generic and therefore can be replicated in other medical training programs. The main limitation of this approach is its scalability when a high amount of videos needs to be tagged. However, in specialized medical procedure training, as CVC installation, it is a viable option. Another potential limitation is when the medical procedure has more complex patterns, e.g., when the order among some activities is not relevant. In such a case, more advanced conformance checking techniques should be considered.

In the future, it is necessary to measure the actual impact of feedback on learning. To this end, a feedback methodology will be applied that includes the tagging of videos by the own students, and the delivery of a report based on the results of the process mining analysis after the PRE evaluation. To measure whether the process-oriented feedback has a statistically significant impact on students' learning, their performance on the PRE and POST scenarios will be evaluated using GRS, comparing an experimental group with a control group.

Acknowledgments. This work is partially supported by CONICYT FONDECYT 181162, CONICYT FONDECYT 11170092, CONICYT REDI 170136, VRI-UC Interdisciplinary 2017, and FOND-DCC 2017-0001. We thank Jerome Geyer-Klingeberg and Celonis Academic Alliance for their support and material.

References

1. van der Aalst, W.M.P.: Process Mining - Data Science in Action. Springer (2016)
2. Archer, J.C.: State of the science in health professional education: effective feedback. Medical education 44(1), 101–108 (2010)

3. Bould, M., Crabtree, N., Naik, V.: Assessment of procedural skills in anaesthesia. British journal of anaesthesia 103(4), 472–483 (2009)
4. Chowdhury, R.R., Kalu, G.: Learning to give feedback in medical education. The Obstetrician & Gynaecologist 6(4), 243–247 (2004)
5. Cook, D.A., Hatala, R., Brydges, R., Zendejas, B., Szostek, J.H., Wang, A.T., Erwin, P.J., Hamstra, S.J.: Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. Jama 306(9), 978–988 (2011)
6. Corvetto, M., Pedemonte, J., Varas, D., Fuentes, C., Altermatt, F.: Simulation-based training program with deliberate practice for ultrasound-guided jugular central venous catheter placement. Acta Anaesthesiologica Scandinavica 61(9), 1184–1191 (2017)
7. van Eck, M.L., Lu, X., Leemans, S.J., van der Aalst, W.M.P.: Pm2: A process mining project methodology. In: International Conference on Advanced Information Systems Engineering. pp. 297–313. Springer (2015)
8. Fecso, A.B., Szasz, P., Kerezov, G., Grantcharov, T.P.: The effect of technical performance on patient outcomes in surgery: a systematic review. Annals of surgery 265(3), 492–501 (2017)
9. Günther, C.W., van der Aalst, W.M.P.: Fuzzy mining–adaptive process simplification based on multi-perspective metrics. In: International Conference on Business Process Management. pp. 328–343. Springer (2007)
10. Hagedorn, J., Hailpern, J., Karahalios, K.G.: Vcode and vdata: illustrating a new framework for supporting the video annotation workflow. In: Proceedings of the working conference on Advanced visual interfaces. pp. 317–321. ACM (2008)
11. Ilgen, J.S., Ma, I.W., Hatala, R., Cook, D.A.: A systematic review of validity evidence for checklists versus global rating scales in simulation-based assessment. Medical education 49(2), 161–173 (2015)
12. Lalys, F., Jannin, P.: Surgical process modelling: a review. International journal of computer assisted radiology and surgery 9(3), 495–511 (2014)
13. Leape, L.L., Brennan, T.A., Laird, N., Lawthers, A.G., Localio, A.R., Barnes, B.A., Hebert, L., Newhouse, J.P., Weiler, P.C., Hiatt, H.: The nature of adverse events in hospitalized patients: results of the harvard medical practice study ii. New England journal of medicine 324(6), 377–384 (1991)
14. Ma, I.W., Zalunardo, N., Pachev, G., Beran, T., Brown, M., Hatala, R., McLaughlin, K.: Comparing the use of global rating scale with checklists for the assessment of central venous catheterization skills using simulation. Advances in health sciences education 17(4), 457–470 (2012)
15. Nesbitt, C.I., Phillips, A.W., Searle, R.F., Stansby, G.: Randomized trial to assess the effect of supervised and unsupervised video feedback on teaching practical skills. Journal of surgical education 72(4), 697–703 (2015)
16. Neumuth, D., Loebe, F., Herre, H., Neumuth, T.: Modeling surgical processes: A four-level translational approach. AI in medicine 51(3), 147–161 (2011)
17. Rojas, E., Munoz-Gama, J., Sepúlveda, M., Capurro, D.: Process mining in healthcare: A literature review. Journal of biomedical informatics 61, 224–236 (2016)
18. Rojas, E., Sepúlveda, M., Munoz-Gama, J., Capurro, D., Traver, V., Fernandez-Llatas, C.: Question-driven methodology for analyzing emergency room processes using process mining. Applied Sciences 7(3), 302 (2017)
19. Walter, A.J.: Surgical education for the twenty-first century: beyond the apprentice model. Obstetrics and Gynecology Clinics 33(2), 233–236 (2006)
20. Zendejas, B., Wang, A.T., Brydges, R., Hamstra, S.J., Cook, D.A.: Cost: the missing outcome in simulation-based medical education research: a systematic review. Surgery 153(2), 160–176 (2013)