Medical Training Analytics through Process Mining: The Case of Central Venous Catheters Surgeries

Conformance Checking Challenge 2019*

Jerome Geyer-Klingeberg, Janina Nakladal,
Peyman Badakhshan, Iyappan Somasundaram

Celonis SE, Theresienstraße 6, 80333 Munich, Germany
{j.geyerklingeberg, j.nakladal, p.badakhshan, i.somasundaram}@celonis.com

Abstract. The 2019 Conformance Checking Challenge (CCC2019) refers to the medical training process of installing Central Venous Catheters (CVC) with ultrasound, which is performed in two rounds by medical students. This report provides insights into the surgery training performance of the students. Using the Celonis Intelligent Business Cloud (IBC), we discover the particulars of the given event log and derive deviations between the expected process model and the actual procedural behavior for a successful surgery. Various process mining methods are combined in this report, such as conformance checking, process and variant discovery, throughput time and rework analysis, as well as benchmarking of different students and rounds. Interpretable dashboards are provided for the two user perspectives: students and instructors. They allow quick insights into individual performance, the deduction of violation patterns on process stage and activity level, as well as best-practice examples and recommendations for training improvement. The report also discusses limitations due to process and data characteristics, sketches possibilities for additional data collection and improvement of the medical training through violation scoring, and outlines ideas for the creation of actionable real-time recommendations for the students.

Keywords: Process Mining, Conformance Checking, Medical Training, Healthcare, Central Venous Catheters, CCC19, Learning Analytics, Celonis.

1 Introduction

The global shortage of well-trained medical staff is a major threat for our health systems to provide fast, reliable and helpful medical services [1,2,3]. Driven by an increasing and aging population, the demand for healthcare services is rapidly growing, while supply is stagnating. The result is a dramatic demand-supply imbalance of medical workers. The World Health Organization (WHO) estimates that there is a lack of more than 4 million doctors, nurses and other health workers to meet the worldwide demand [4].

---

* Corresponding author: j.geyerklingeberg@celonis.com

Many solutions have been proposed to address the healthcare workforce crisis [2]. One of the most intensively discussed measures is to increase the number of medical graduates by opening new medical schools, growing enrollments numbers in medical study programs, and enlarging classroom sizes [2]. However, the solution itself comes again with problems. Budgets and training staff are often not increasing as fast as enrollments rates and classroom sizes [5]. Thus, teachers are facing the challenge to train more students, while having less time for the individual trainee.

Traditionally, the practical training part in medical education, particularly for surgery training, pursues the "See One, Do One, Teach One" approach [6]. Following this method, medical trainees are capable of performing a specific procedure after observing it once and to teach it to other trainees afterward [7]. However, many concerns have been raised against this learning method [8,9,10]. Especially, patient safety is at risk with this approach. Studies show that 28-42% of trainees feel inadequately trained to perform a practical medical intervention alone for the first time [11,12]. Due to this caveat, modern medical exercises are often conducted in lab-based procedure training. This offers protected training in a "mistake forgiving" environment [13], where the students can practice procedures on dummies, standardized patients or with each other before executing a skill on real patients [14,15,16]. Lab and simulation-based training have been shown to significantly improve procedural skills in medical education [17,18,19].

According to [17], educational feedback is a crucial factor for the learning success of lab and simulation-based training by providing a chance for reflection on the procedural performance. Many studies underline that significant learning progress occurs during debriefing periods following each simulation [20,21]. Earlier, this feedback was mainly based on observations of the supervisors and the evaluation of the outcome. However, the rapidly changing technology, like virtual patients, e-learning or video tracking methods, create the opportunity for data-based feedback [22]. Therefore, the trainers can take advantage of the data gathered during the lab exercises. However, to translate this data into valuable insights for personal and target-oriented feedback, powerful analytics are required to evaluate the training data. Process mining is a promising technique to enable such a data-based learning analysis. It has successfully been applied to derive data-based feedback in online learning services [23,24], but also in surgical training [25].

Process mining is a disruptive data-analytics technology that is used to discover, analyze, and predict real processes by extracting knowledge from event log data that is readily available from modern information systems [26]. The starting point of process mining is the event log. Each event consists of defined steps in a process (activities), information about the time of the activity's execution (timestamp), and the related process instance (case). The activities of each case are ordered according to the timestamp, which allows illustrating the flow of the real process. Additional information, as for example the person or device executing an activity, are often added to extend the event log by context-related information. As illustrated in *Figure 1*, three types of process mining are commonly applied [26]: the reproduction of process models based on the event log data (discovery), the comparison of these process models with prescribed behavior (conformance), and the extension/improvement of process models using the information from discovery and conformance (enhancement).

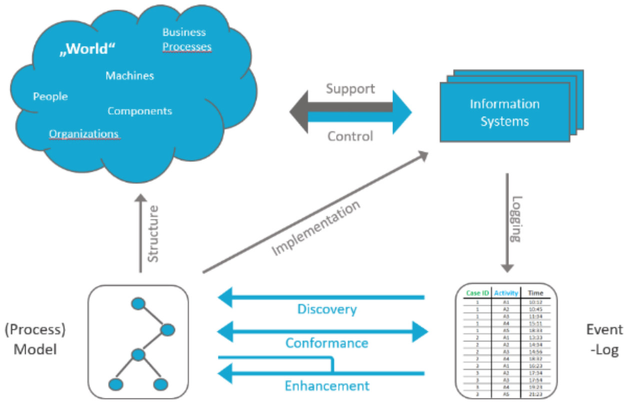


Fig. 1. Process mining overview

The advantage of process mining lies in its process-oriented view. This is especially helpful for the training of medical procedures. For example, surgery can be described as a process starting with the patient preparation, followed by the steps of the surgery itself, and a defined end, e.g. the closing of the incision. Usually, there is a clear structure describing the order of the different steps within this process.

In this paper, we apply process mining for the analysis of the Central Venous Catheter (CVC) installing process[1]. The data is provided by the Conformance Checking Challenge 2019 (CCC19) [27]. We present and apply a novel analytic framework based on conformance checking to compare the actual training performance of the students with the prescribed CVC process in order to provide personal and tailored training feedback. In the analysis, we are taking in two perspectives: the instructor and the student. In addition, we are looking at process stages and the activities within the stages to gain both high-level and detailed insights.

The findings are summarized in easy-to-use analysis dashboards in the process mining software Celonis. The key results of this analysis are reported as an overview for the instructors and students in *Table 1*.

---

[1] The CVC transports medicine, nutrition or fluid in the body. In contrast to a standard catheter, the CVC remains in the central vein. Thus, installing it correctly is critical. See also, https://www.thoracic.org/patients/patient-resources/resources/central-venous-catheter.pdf.

Table 1. Summary of results

Students	Instructors
• 1 out of 10 cases conforming on a stage-level in the PRE stage. • 0 out of 10 cases conforming on an activity-level in the PRE stage. • Major deviations in terms of throughput-time and number of activities, resulting in a heterogenous training group. • Possible deep dive into single process flow through Process and Variant Explorer. • Focus of exercise sessions needs to be on unscrambling of stages and activities to avoid rework and critical incidents.	• Activities become more streamlined from PRE to POST • Rework appears more often in the critical later stages of the process when installing the guidewire and catheter, as well as with gateways • Average throughput times decrease from 47 minutes (PRE) to 17 minutes (POST). • Conformance rates and number of violations improve from PRE to POST round. • On the stage level, 3 conforming cases can be found across both rounds.

These insights support both the instructors and the students. Instructors can easily monitor the performance of each student and check the deviation between their behavior and the ideal process of installing CVC to give tailored and evidence-based feedback. Students receive valuable insights to understand their issues in conducting the process and improve it for future actions.

The remainder of this paper is structured as follows. Section 2 describes the data. The analysis approach is outlined in Section 3, followed by the results for the students' perspective are reported in Section 4, whereas Section 5 covers the results for the instructors' perspective. Section 6 includes recommendations for future work, and Section 7 concludes.

2 Data Overview

In this section, we present our understanding of the provided data and the process. This understanding is the base for the analysis and interpretation of the process in Section 4 and 5.

2.1 Event collection

The data for the CC Challenge 2019 has been collected within the interdisciplinary research project titled "Process-Oriented Medical Education (POME)" conducted by the School of Medicine and the School of Engineering of the Pontificia Universidad Católica de Chile (PUC) [27]. The School of Medicine at PUC developed a simulation-based training for the ultrasound-guided jugular central venous catheter placement based on four stages [25,28]:

1. *Instructions*: The trainers teach the students how to perform the procedure
2. *PRE recording*: The students execute a first procedure of the Central Venous Catheter (CVC) installation to benchmark their skills. Videos are recorded individually for each student.

3. *Practice*: The trainees complete multiple exercise sessions at their discretion.
4. *POST recording*: Another video is recorded for the second execution of the CVC installation to evaluate the skills they acquired during the training program.

The students conduct the two exercises of the surgery (PRE and POST) on a mannequin. Although this simulates many aspects of the real surgery, there are less unforeseen actions (e.g. unexpected blood loss, patient dies, etc.).

In contrast to classical event collection in process mining, the event log data is not automatically tracked during the process execution. Instead, two medical experts watched the recording of each student using special software for video tagging. The tags are set for each activity, such that an event log is generated including the activity name, the start and end time of each activity.

2.2 Event log

The event log generated by tagging the video recording of the PRE and POST training executions of the CVC installation includes 20 cases (10 PRE and 10 POST executions) for 10 different students, 697 events and 29 different activities. The events have been executed in the period between October 2018 and January 2019. *Table 2* represents the attributes of the event logs and their descriptions.

Table 2. Overview of the data attributes

Attribute	Description
CASEID	ID (unique identifier) for each training trace
RESOURCE	Student ID executing the training procedure
ROUND	PRE (if the video was recorded before the practice session) POST (if the video was recorded after the practice session)
EVENTID	ID of the event
ACTIVITY	Name of the executed activity
STAGE	Stage of the CVC procedure (each activity is assigned to a specific stage)
START	Time when the activity started
END	Time when the activity ended
VIDEOSTART	Time in the video recording when the activity started
VIDEOEND	Time in the video recording when the activity ended

All timestamps are reported by the exact date (mm-dd-yyyy) and time (hh:mm:ss). This dataset is suitable to perform all three types of process mining (*Figure 1*).

2.3 Understanding of the Process

Performing the surgical procedure of the CVC installation can be described as a process covering a set of activities executed in a specific order. Activities can be grouped within six stages (sub-processes) of the surgery. *Table 3* outlines the 29 different activities performed during the PRE and POST training, to which stage each activity refers to, as well as the number of occurrences in the full event log (N_{Total}), the subsample of the PRE (N_{PRE}) and POST phase (N_{POST}). The last column reports the difference in the number of occurrences between POST and PRE to see how the execution changed after the training. Negative values indicate that an activity is executed less frequently after the training, positive values show that the execution rate increased.

Table 3. Process stages and activities

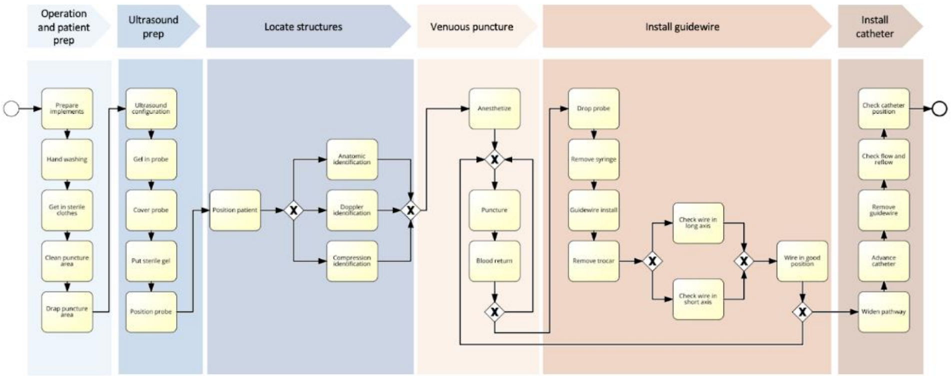
Stage	Activity	N_{Total}	N_{PRE}	N_{POST}	N_{POST}-N_{PRE}
1. Operator and patient preparation	Prepare implements	63	29	34	+5
	Hand washing	22	12	10	-2
	Get in sterile clothes	37	18	19	+1
	Clean puncture area	19	9	10	+1
	Drap puncture area	20	10	10	+/-0
2. Ultrasound preparation	Ultrasound configuration	33	18	15	-3
	Gel in probe	19	9	10	+1
	Cover probe	20	10	10	+/-0
	Put sterile gel	22	12	10	-2
	Position probe	25	14	11	-3
3. Locate structures	Position patient	6	4	2	-2
	Anatomic identification	25	15	10	-5
	Doppler identification	5	2	3	+1
	Compression identification	21	11	10	-1
4. Venous puncture	Anesthetize	19	9	10	+1
	Puncture	30	19	11	-8
	Blood return	29	18	11	-7
5. Install guidewire	Drop probe	28	17	11	-6
	Remove syringe	26	15	11	-4
	Guidewire install	28	17	11	-6
	Remove trocar	24	13	11	-2
	Check wire in long axis	31	18	13	-5
	Check wire in short axis	27	15	12	-3
	Wire in good position	10	2	8	+6
6. Install catheter	Widen pathway	21	12	9	-3
	Advance catheter	30	15	15	+/-0
	Remove guidewire	21	11	10	-1
	Check flow and reflow	19	9	10	+1
	Check catheter position	17	7	10	+3

In the subsequent analysis, we differentiate between the *high-level* (stage-based) and the *detailed-level* (activity-based) analysis. As the timestamps in the event log refer to the start and end time of each activity, we define the start time of the first activity within a stage as stage starting time and the end time of the last activity within this stage as stage end time.

2.4 The Process Model

As opposed to the event log including the real process data, the process model defines a reference model describing the prescribed behavior in the process [29]. The model for the CVC installation was defined using a Delphi method [27]. A panel of 13 medical experts from three specialty areas and eight different medical institutions were asked to define the CVC process. Based on this input, the process model was designed to show which activities should be conducted in which order. *Figure 2* illustrates the process model for the CVC surgery in the BPMN (Business Process Modelling Notation). The model reflects both the *stage-based* view (indicated by colors) and the *activity-based* view (indicated by rounded-corner rectangles).

Fig. 2. Process model for CVC installation



To allow decisions in the process in the sense of either/or logic, the BPMN model includes exclusive gateways. In each decision point, which is marked by a diamond shape form including a ‘x’, only one flow can be taken. In the analysis, we use the process model to compare it with the actual process data from the event log and explore their deviations (conformance checking).

2.5 Two Process Perspectives

There are two groups of persons to be considered as recipients of the process mining insights from the CVC training data: (i) instructors, and (ii) students. Both stakeholders are characterized by different expectations towards the focus of the analysis and the questions to be answered as summarized in *Table 4* [27]. The two perspectives and their requirements are also considered in the following analysis.

Table 4. Instructors’ vs. students’ perspective

Perspective	Aggregation level	Focus	Desired insights	Goal
Instructors	High	Class performance	Common mistakes, similarities/differences across groups	Improve training concept, identify need for additional training
Students	Low	Individual performance	Personal mistakes, especially after PRE round	Improve performance in POST test

2.6 Data Preparation

As the event log is created by manual tagging of the activities recorded in the training sessions, there is a certain subjectivity included, which may lead to noise or wrong assignment of activities and time stamps. Therefore, the event log has first been checked for any misleading entries or missing data. For the user R_13_C in the POST training, we find that there is an error in the timestamps between Position probe (end time 01:16:2019 1:09.40) and the subsequent activities (next activity start time 01:17:2019), which starts one day later. We account for this error by adjusting the date of these activities from 01:17:2019 to 01:06:2019.

After that, the event log file (CSV format) was imported in the Celonis Intelligent Business Cloud (IBC) to create and configure the data models used as basis for the analyses. As *case* identifier, the column CASEID was used, describing the ID of the video for one student in one particular round. To cover both, the stage-based view and the activity-based view in the analysis, two different data models were created. One contains the stages (STAGE) as *process activities*, enabling a high-level view on the process and a general understanding of its structure. Second, the ACTIVITY column is used to describe process activities. This second data model leads to more granular insights based on every single process activity. The data models are the input for the subsequent analysis.

The event log contains the start (START) and end time (END) of activities and additionally the start (VIDEOSTART) and end time (VIDEOEND) of activities in the video recording (see *Table 2*). For setting up the data model in Celonis, the START and END were used as timestamps. This allows analyzing both: the throughput time of each activity, but also the time of the transition between two consecutive activities. Additionally, the video start time was used as a sorting logic to link it to the timestamps and to have a sorting if two activities have identical start or end time.

3 Analysis structure

In this section, we outline the analysis approach. The findings derived from this analysis are presented in the subsequent Section 4.

3.1 The analysis tool

The data preparation and all analyses have been performed in the Celonis IBC, which is the market-leading process mining tool [30]. Celonis provides a very fast algorithm to visualize process models and analysis components, which are easy to interpret for the interdisciplinary audience of this challenge. Moreover, Celonis offers a very clear user interface that guarantees high usability for both instructors and students and a fast understanding of the actual training performance and needs for improvement.

3.2 Filtering

Celonis integrates various filters and selections to drill down the process interactively. We use two types of filters:

- *Case filter*: We use this filter to analyze the activities performed by a specific student and to differentiate between activities conducted in the PRE or POST stage.
- *Activity filter*: This filter allows to select certain activities or to exclude them from the analysis. For example, we compare the process models flowing or not flowing through certain important activities.

3.3 Performance measures

In Celonis, we define certain key performance indicators (KPIs) using the natively integrated Process Query Language (PQL). Especially, we focus on three main KPIs:

- *Throughput time*: Refers to the time to conduct a specific process step (activity time) or the time between certain process steps (transition time) in minutes. The overall throughput time illustrates the time between process start and process end.
- *Number of activities*: Refers to the number of occurrences of the same activity (rework activities) or the number of activities per case, i.e. per student and round (total number of activities).
- *Conformance rate*: Refers to the percentage rate of cases which are conforming with the prescribed target model (*Figure 2*).

3.4 Methods

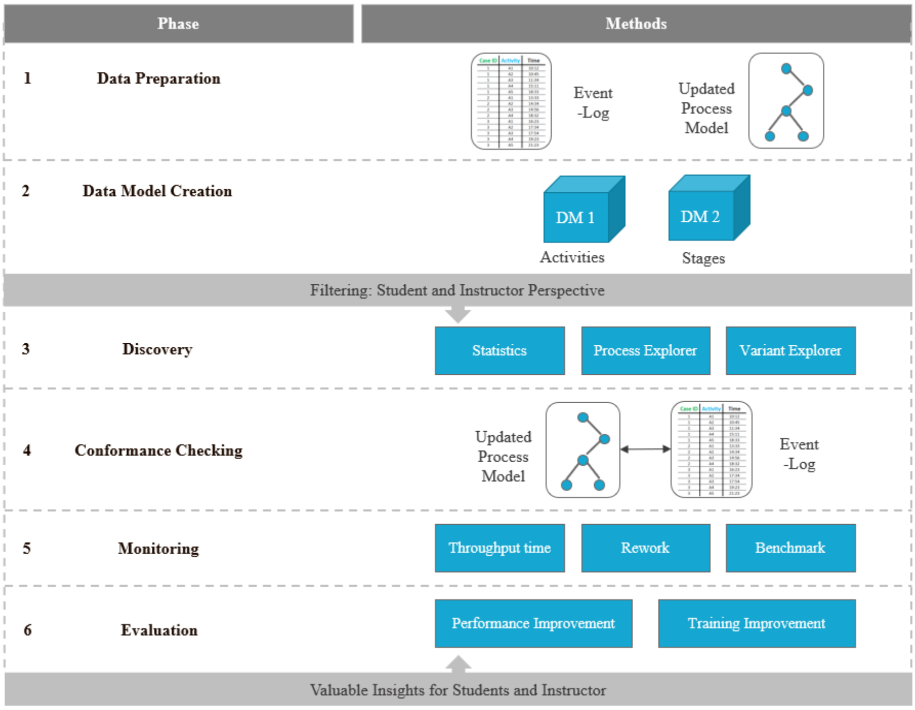
We use the following process mining methods for our analysis:

- Process Discovery: Variant and Process Explorer
- Conformance Checking: Conformance Checker and Social Network Analysis from the Celonis Proactive Insights (PI) features, including a root cause analysis feature to detect the causes for deviations of the process data from the target process model [31]
- Enhancement: Various process analytics components, especially graphics (like histograms, charts or boxplots) and tables (like OLAP)

3.5 Analysis framework

To derive the findings for this report, we follow a 6-step procedure, leading from data preparation to the evaluation of results for both instructors and students. This approach is outlined in *Figure 3*.

Fig. 3. Analysis framework



After preparing the data and generating the two data models (see Section 2.6), we start with discovering the actual process flow and its different variants, using the Process and Variant Explorer in Celonis. Several tables with process details and measures for the three main KPIs are used to extend the insights from the graphical process flow analysis. In the next step, the process model (*Figure 2*) is used as a reference model to be compared with the actual data with the help of the Conformance Checker.

Knowing the deviations on an activity-base, several analyses were created in the monitoring stage to enhance the understanding of the actual training performance. They include a benchmarking of students and rounds and a cycle time analysis, to focus on deviations in terms of timing, as well as a PI Social analysis.

Finally, the results were compiled and interpreted. Starting with the instructor and the high-level of aggregation and following with individual students' performance, this paper focuses on a broad overview and explanation of deviations and actual performance. For the individual students, some examples are highlighted, without describing each case. The objective of this analysis framework is to provide each stakeholder with meaningful and broad insights into the medial training process, that are easy to understand and interpret.

In the next section, we report our results derived through the analysis structure described in the previous section. We follow a certain storyline, which demonstrates the usability of the insights for the two stakeholders: Exemplary students are looking at the data after the PRE round and find out about their mistakes in terms of stages/activities, their order and timing. They use the insights for the practice sessions to prepare for the final POST test. After both rounds, the instructor analyzes the full data set and evaluates the actual performance of the group. The instructor can compile findings, benchmark certain students/rounds, and analyze cycle times to understand the impact of the training on the performance of students.

4 Results for Student's Perspective (PRE)

In order to improve their procedural activities, the students need an overview of their conformance with the process model derived from the expert panel and the violations to be improved. The analysis starts on a high level with the stage-based data model and continues with insights into the activity-based performance of three exemplary students.

4.1 Conformance Checking

Aligning the event data and the reference model on the stages-level results in 3 conforming cases against 17 non-conforming cases. Moreover, 23 violations can be detected across both rounds (PRE and POST). Filtering on the PRE round results only, we find 1 conforming case against 9 non-conforming cases and 21 violations.

Looking at the list of violations, several of them refer to "followed by" problems. As each stage consists of several activities, all deviations related to one stage being followed by the same stage, e.g. Locate Structures is followed by Locate Structures, are not necessarily a process violation. For the sake of having a clear overview of critical incidents, the 6 deviations of stages followed by the same stage are whitelisted. This had no effect on the conformance rates and KPIs.

The conformance is checked for each student individually, so that everyone has an overview of the respective performance in the PRE round. *Table 5* provides an overview of the stage-based conformance checking and the key KPIs (on activity-level) for each student (CR = Conformance Rate of stages in %, #V = Number of violations in stages (including whitelisted violations), TPT = Total throughput time from training start to training end, AT = Total time for activity execution, TT = Total transition time for all transitions between activities, #A = Total number of activities, #R = Number of activities performed more than once).

Conformance rates differ between 0% and 100%, violations in total range between 6 and 13 (including the whitelist). Total throughput times vary between 15.8 minutes and 93.7 minutes. Looking at the transition time, there occur peaks up to 68 minutes between two activities, resulting also in a large total throughput time. The number of activities in the process differs between 26 and 59, the number of reworking activities differs between 1 and 14. The complexity of the underlying medical training process can also be seen in the varying KPIs across students. In addition, the results point towards a heterogenous group of students.

Table 5. Conformance results and KPIs for each student

Student	CR	#V	Violations (not whitelisted)	TPT	AT	TT	#A	#R
R_13_1C	0	8	• Locate Structure - Ultrasound Preparation • Ultrasound Preparation - Operator and Patient Preparation • Ultrasound Preparation - Venous Puncture	22 min	15 min	8 min	33	6
R_14_1D	0	10	• Locate Structure - Operator and Patient Preparation • Operator and Patient Preparation - Locate Structures • Ultrasound Preparation - Venous Puncture • Venous Puncture - Ultrasound Preparation	21 min	17 min	1 min	31	5
R_21_1F	0	13	• Install Catheter - Install Guidewire • Install Catheter - Ultrasound Preparation • Install Guidewire - Ultrasound Preparation • Install Guidewire - Venous Puncture • Locate Structures - Operator and Patient Preparation • Ultrasound Preparation - Operator and Patient Preparation • Ultrasound Preparation - Venous Puncture	94 min	27 min	68 min	59	14
R_31_1G	0	12	• Install Catheter - Ultrasound Preparation • Locate Structures - Operator and Patient Preparation • Locate Structure - Ultrasound Preparation • Operator and Patient Preparation - Venous Puncture • Ultrasound Preparation - Operator and Patient Preparation • Venous Puncture - Operator and Patient Preparation	90 min	25 min	67 min	46	14
R_32_1H	0	9	• Install Guidewire - Venous Puncture • Locate Structures - Ultrasound Preparation • Ultrasound Preparation - Operator and Patient Preparation	82 min	17 min	66 min	36	8

R_33_1L	0	10	• Install Guidewire - Venous Puncture • Locate Structures - Operator and Patient Preparation • Locate Structures - Ultrasound Preparation • Operator and Patient Preparation - Locate Structures • Ultrasound Preparation - Operator and Patient Preparation	26 min	18 min	10 min	35	7
R_45_2A	0	8	• Install Guidewire - Venous Puncture • Ultrasound Preparation as START Activity	21 min	17 min	6 min	38	10
R_46_2B	0	9	• Locate Structures - Ultrasound Preparation • Ultrasound Preparation - Operator and Patient Preparation • Locate Structures as START Activity	16 min	13 min	3 min	30	3
R_47_2C	100	6	• All whitelisted	23 min	18 min	5 min	26	1
R_48_2D	0	8	• Install Catheter - Venous Puncture • Locate Structures - Ultrasound Preparation • Ultrasound Preparation - Operator and Patient Preparation	77 min	11 min	67 min	36	11

Only one student (R_47_2C) has 100% conformance in terms of stages and in addition a very short throughput time (23 minutes), resulting in 26 activities being performed. This student can be classified as a benchmark within the group. In contrast, another student (R_21_F) has 13 violations, whereas 7 are critical and not whitelisted. The student needs 93.7 minutes to conduct the procedure and performs 59 activities in total. It needs to be checked in the next section, whether the throughput time is caused by one single transition or the amount of activities performed. Another interesting case is student R_46_2B, who performs the Location of Structures as the start activity of the process without preparing the patient and ultrasound. However, the student is the quickest in terms of throughput time. This shows that the throughput time itself is not a single dimension that should be minimized but must instead be evaluated in combination with the conformance rate. Ideally, we find high conformance and low throughput times.

Besides the stage-based view, students can also check the conformance in the activity-based process model. Continuing with the three exemplary students, the conformance checker based on activities, filtered on PRE round and each student unveils the violations in terms of activities, whereas every student has a conformance rate of 0% as compared to the process model proposed by the expert panel.

R_47_2C, with a perfect stage-conformance, shows 7 violations, mainly related to a different order of activities in the first two stages, which are not too critical. A more severe deviation occurs: After removing the trocar, the pathway is widened, leaving out the checking of the wire. R_46_2B started with an incorrect stage, which can also be seen in the list of 8 violations based on deviating activities: Position the patient is conducted as a start activity. For R_21_F, with a very long throughput time, the activity-based conformance check displays 19 violations, which can mainly be clustered in jumping back to prior activities at later stages: Removing the guidewire is followed by positioning the probe, position probe is followed by puncture and others. This leads to the conclusion that this student might be insecure or especially concentrated on quality, continuing with re-checking activities.

Lastly, the underlying process model does not include any parallel gateways (i.e. tasks should not be executed in parallel). Nevertheless, some students start activities in parallel, especially in the first stage of the process. One example is student R_46_2B, that gets in sterile clothes and puts gel in probe at the same time. While parallel execution of activities can improve throughput time, it might not be a desirable approach in medical training, as students shall focus on the correct step-by-step procedure of tasks for a high-quality execution of the CVC installation.

4.2 Process Discovery

The conformance checker provides students with an overview of the conformance rate in terms of stages/activities and the key KPIs. A closer look at the process flow of each student via the Process and Variant Explorer unveils the exact process order of the PRE round and problems within each stage or certain activities. It helps to find loops and rework within the procedure. For the activity-based data model, the activities have a color format according to their stages for a quick grouping and easier analysis (see *Figure 4*).

For R_21_F, the Process Explorer on stage-base shows that the long throughput time mainly occurs in the transition from Installing the guidewire to Venous puncture, where it takes 60 minutes (*Figure 5*). Having a closer look at the activity-level, it becomes apparent that the student performs the activities fast, but there is one outlier in the transition between drop probe and puncture, where the student stops for 60 minutes. As all other transitions take 1 minute or shorter, the student might have had a break in the procedure.

For R_47_C, the process flow is quite streamlined and at a 100% conformance for the stage-level. On an activity level, the order is slightly different for the first activities. Student R_46_2B has problems with the start of the process which can easily be seen in the colored Process Explorer: Position the patient is the start activity, from stage 3, followed by ultrasound configuration from stage 2, followed by getting into sterile clothes. The student probably started the wrong way and tried to get back to the right procedure.

Fig. 4. Coloring on stage-level to improve usability in activity-based Process Explorer

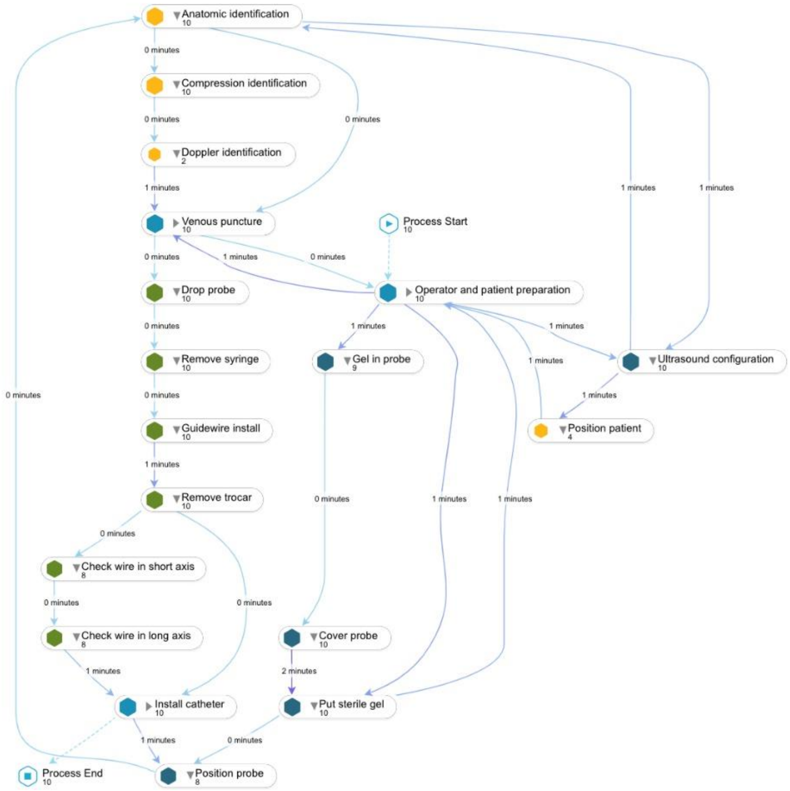
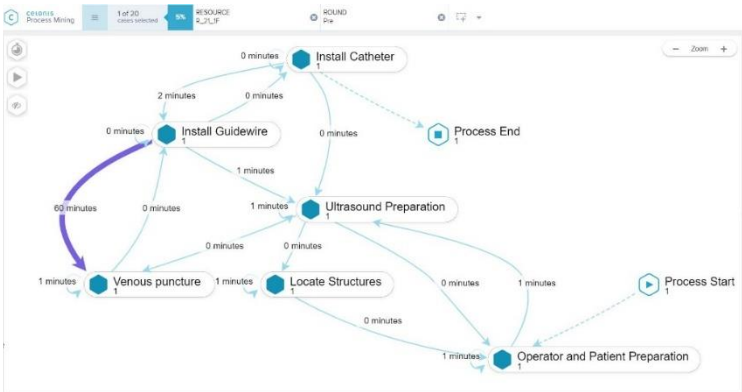


Fig. 5. Stage-based Process Explorer view for student R_21_F



Overall, the performance of students after the PRE round can be evaluated based on stages and activities. Using the conformance checker, filtering, and the Process and Variant explorer in Celonis, every student gets a personal dashboard providing a clear picture of the performed CVC procedure evaluated by the three KPIs. They can use these insights to work on the overall order and timing of stages, but also on single activities, as well as the throughput time in terms of single activities and transitions. Providing students with a clear overview and the possibility to dive deeper into the process flow ensures easy to use results and a successful improvement in the POST round.

5 Results for Instructors' Perspective (PRE and POST)

While the students strive to get quick insights into the right-or-wrong status of training procedure and into the actual executed events to improve their performance in the POST test, instructors need a bigger picture to understand the whole training group, their need for additional explanations, as well as the impact of certain training measures. Therefore, instructors need certain KPIs to measure training success, and the possibility to dive deeper into process flows if necessary.

The following KPIs are critical to get a comprehensive overview of the training performance: throughput time, rework activities, and conformance rate. A cross-section emerges from the benchmark between PRE and POST round.

For the instructor, the reporting of results for throughput time and rework is limited to the activity-level for a preferably granular view to derive meaningful implications. For the conformance rate, there will be a short comparison between stage and activity level for a comprehensive view. All analyses can also be conducted on the stage-level.

5.1 Throughput time

To get a detailed picture of the time required by each student to perform the procedure, we analyze the total throughput time between the training start and the training end (TPT). We further decompose into the time for the execution of the activities (AT) and the time for the transition between the activities (TT). The results are reported in *Table 6* for each student (detailed view) and corresponding summary statistics can be found in *Table 7*.

Table 6. Overview of cycle times per students

	Total throughput time in min (TPT)			**Total activity time in min (AT)**			**Total transition time in min (TT)**		
Student	PRE	POST	POST - PRE	PRE	POST	POST - PRE	PRE	POST	POST - PRE
R_13_1C	22	15	-7	15	11	-4	9	3	-6
R_14_1D	21	17	-4	17	15	-2	8	3	-5
R_21_1F	94	16	-78	25	13	-12	68	4	-64

R_31_1G	90	16	-74	25	14	-11	67	2	-65
R_32_1H	82	21	-61	17	16	-1	66	7	-59
R_33_1L	26	13	-13	18	11	-7	10	2	-8
R_45_2A	22	18	-4	17	15	-2	6	5	-1
R_46_2B	16	13	-3	13	11	-2	3	3	+/-0
R_47_2C	23	11	-12	18	8	-10	5	3	-2
R_48_2D	76	25	-51	11	20	9	67	5	-62

Table 7. Summary statistics of cycle times

	Total throughput time in min			**Total activity time in min**			**Total transition time in min**		
	PRE	POST	POST - PRE	PRE	POST	POST - PRE	PRE	POST	POST - PRE
Minimum	16	11	-5	11	8	-3	3	2	-1
Maximum	94	25	-69	25	20	-5	68	7	-61
Mean	47	17	-31	18	13	-4	31	4	-27
Median	25	16	-9	17	14	-4	10	3	-7
Std. Dev.	33	4	-29	5	3	-1	31	2	-30
Mean diff. test	p-value: 0.01			p-value: 0.06			p-value: 0.02		

We see that all throughput times declined in the POST round. However, especially the transition time reduced by 27 minutes on average, as compared to a reduction of 4 minutes on average in the activity time. This indicates that the students work much more efficient, with less time to think and prepare between the activities. These results are confirmed by a mean difference test (paired *t*-test for non-independent samples) between the PRE and the POST round. The null hypothesis of the mean values being equal in the PRE and POST round can be rejected at the common level of 1% (highly significant) for the total time and the transition time at the common level of 5% (significant). For the activity time, we find no statistically significant result to reject the null hypothesis.

5.2 Rework activities

Reporting the throughput time of training procedures is strongly linked to the performed activities and possible rework activities, meaning loops or repeatedly performed activities. *Table 8* provides an overview of the number of occurrences for each activity (PRE and POST) to quickly identify rework. Empty fields indicate that this activity was not conducted.

First, looking at the activities and counting their occurrence in total and per case, it becomes apparent that not all activities are performed by every student and some activities are performed significantly more often. While some of the less performed activities are part of a gateway path and optional, others are mandatory but not conducted by every student: Position patient, Wire in good position, Check catheter position, Anesthetize, Position probe, Widen pathway, Check flow and reflow, Clean puncture area and Gel in probe. It needs to be checked if students understood their necessity in the previous training or if more emphasis is required in the future.

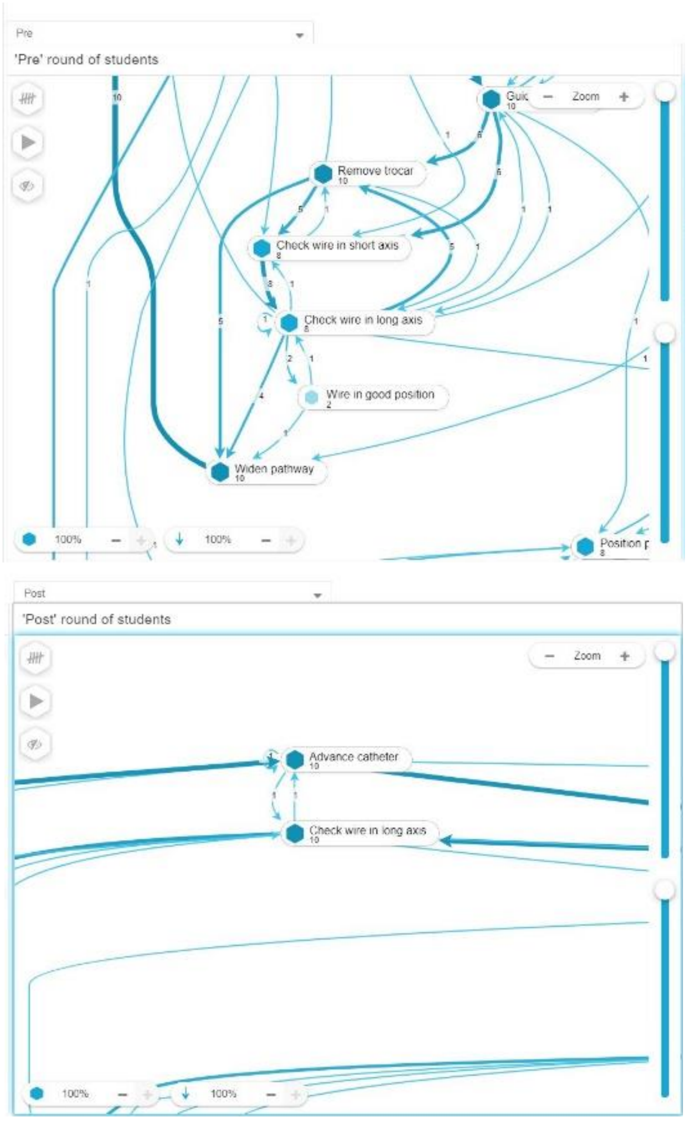
Table 8. Rework overview

Activity	PRE (No. of occurrences)					POST (No. of occurrences)				
	0	**1**	**2**	**3**	**>3**	**0**	**1**	**2**	**3**	**>3**
Advance catheter		6	3	1			9			1
Anatomic identification		7	1	2		1	8	1		
Anesthetize	2	7	1				1			
Blood return		5	3	1	1		9	1		
Check catheter position	3	7					1			
Check flow and reflow	1	9					1			
Check wire in long axis	2	4	2		2		7	3		
Check wire in short axis	2	5	1	1	1	1	7	1	1	
Clean puncture area	1	9					1			
Compression identification		9	1			1	8	1		
Cover probe		1					1			
Doppler identification	8	2				7	3			
Drap puncture area		1					1			
Drop probe		5	4		1		9	1		
Gel in probe	1	9					1			
Get in sterile clothes		2	8				1	9		
Guidewire install		4	5	1			9	1		
Hand washing		8	2				1			
Position patient	6	4				8	2			
Position probe	2	4	3		1		9	1		
Prepare implements		2	3	2	3			2	3	5
Puncture		5	3		2		9	1		
Put sterile gel		8	2				1			
Remove guidewire		9	1				1			
Remove syringe		5	5				9	1		
Remove trocar		7	3				9	1		
Ultrasound configuration		3	6	1			5	5		
Widen pathway		8	2			1	9			
Wire in good position	8	2				2	8			
SUM	**36**	**175**	**59**	**9**	**11**	**21**	**230**	**29**	**4**	**6**

The most repeatedly performed activities are Drop probe, Guidewire install, Remove syringe and Remove trocar (all in the same stage), followed by Remove guidewire and Advance catheter. Those activities occur in the later, critical stages of the CVC procedure. Their repeated execution might be related to students checking the important steps again or having to perform difficult activities twice. Both would be important in a medical training process and highlights a vital training area. Lastly, filtering this table per students provides the instructor with a more detailed view on the influence of single students on the number of activities.

In a next step, activities are considered in the Process Explorer, benchmarking the PRE and POST round of students. Displaying all activities and connections, POST round looks slightly more streamlined than PRE round. Zooming in, loops can easily be identified, as well as back and forth connections between activities. The Process Explorer displays the number of times a connection or activity is performed, and the thickness of arrows also depends on this. In addition, the Variant Explorer shows which cases (being students) execute which activity.

Fig. 6. Activity-based Process Explorer for PRE and POST round with loops



Looking at the exemplary *Figure 6*, it comes clear that one loop and rework occurs while checking the wire in the long or short axis. This is a pair of activities in the critical stage of installing the guidewire, so double-checking might be necessary. However, it is important to note, that the process model contains an exclusive gateway at this point, so either the long axis or short axis is checked.

The instructor should therefore also focus on those gateway points, were students should be able to decide on which activity to perform in order to conduct the right procedure on time. Checking how activities flow from one activity to another for each student provides the necessary insights: The connection from removing the trocar to checking the wire in long axis is performed by 15 students (5 students in the PRE round and 10 in the POST round). The path from removing the trocar to checking the wire in short axis is performed by 15 students (6 in PRE and 9 in POST). So nearly all students are checking the wire repeatedly in the POST round to be sure of the procedure.

Teaching students how to make confident decisions in critical situations is important in the healthcare sector and has an impact on rework rate, quality and throughput time of the process.

5.3 Conformance Rate

Chapter 4.1 already provided some insights into the conformance rate of students in the PRE round on a stage-level and activity-level. *Table 9* compiles conformance rates and number of violations aggregated across the training group for each round as well as activity-level vs. stage-level. For the stage level, 6 violations have been whitelisted (see 4.1) (CR = Conformance Rate in %, #V = Number of violations).

Table 9. Conformance across activities and stages

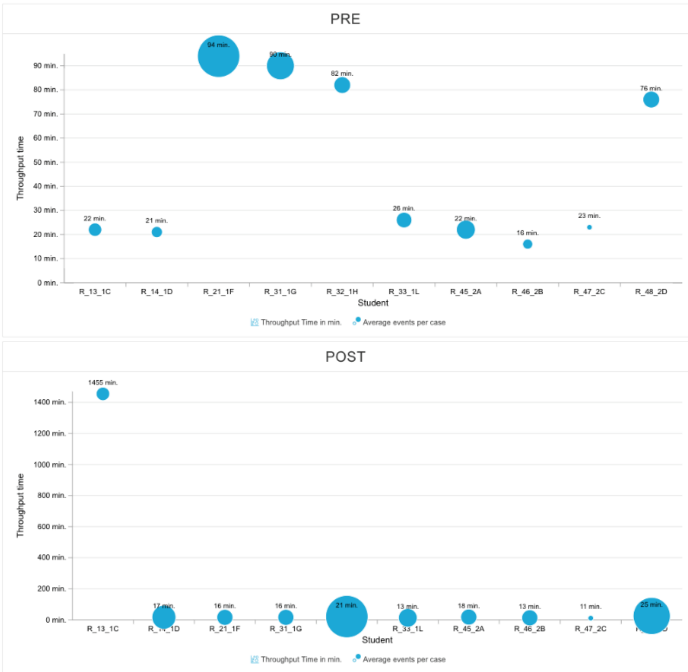
View	Total (CR / #V)	PRE (CR / #V)	POST (CR / #V)
Stage	15 / 23	10 / 21	20 / 17
Activity	0 / 88	0 / 73	0 / 44

The table displays that the conformance rate and number of violations improves from PRE to POST and the performance gets closer to the intended model. However, the conformance rates in general are quite low, highlighting the complexity of the medical training process and also the heterogeneity in the performance of students. This is a clear sign for a need for individual feedback, whether it is data-based or in person. The better and more precise the feedback gets, the more likely will students improve.

The results also show that 3 students across both rounds manage to perform the stages in the right order (one in PRE and two in POST), but no one performs the activities correctly from start to end. The three students on the stage level can be used as a benchmark or best practice example, especially the one that improved from PRE to POST.

Overall, the instructor can use these results for evaluating the whole training group in both rounds and measure training success in terms of the described KPIs. The results are reported in tables for an overview in this paper, but the instructor can access the analyses via easy-to-use dashboards in Celonis, consisting of several components to visualize process flows and report KPI performance. One example is the bubble chart to analyze throughput times (*Figure 7*): It compares both training rounds for each student. The instructor views the group performance but also the single student. The total throughput time is depicted by the position of the bubbles in the chart. In addition, the size of the bubble is linked to the number of average events per case.

Fig. 7. Bubble chart for throughput time in PRE and POST round



Hence, the visualization contains several KPIs, which can also be configured differently (e.g. showing the activity or transition time) and enable the instructor to get as many insights as possible in a short time. This desired approach applies to all other dashboards and analyses as well.

6 Recommendations and future work

Beyond the analysis framework presented in this report, we see various opportunities for improvements and extensions for further analyses.

Data extension. The event log provided for this challenge is sufficient to apply all types of process mining (*Figure 1*). However, the analysis capabilities are restricted by the attributes described in *Table 2*. Therefore, we recommend gathering additional attributes during the event collection to extend the coverage of the analysis. We see the following attributes as valuable additional information:

- Training specific information: type of training (online, offline, type of exercises), ID of teachers, number of training hours, …
- Personal information about students (if compatible with data protection regulation): age, semester, number of training hours, quality rating of the surgery outcome (e.g. judged by one of the experts), …
- Group information (if data is collected across multiple training courses): ID of teacher for this group, …

Especially the more in-depth information about the training procedure could help the instructor to get valuable insights in his analysis. For example, to find out which kind of teaching is the best and most efficient for each activity (practical exercise, video, reading, …).

Automatic event collection. A distinct feature of the data is the way of the event collection via video tagging, which is costly due to the experts to be paid for the tagging and time-consuming leading to a lag between the actual exercise and the availability of the data for analysis). Modern hardware could help to automatically track the event-specific information during the execution of the training. For example, via IoT connected training instruments or RFID tracking [32]. Celonis supports the direct import from data collected through such cyber-physical systems and data generated by (software) robots [33].

Real-time feedback. The automatic event collection is the basis for real-time training feedback. When data about the activities is gathered during its execution, we see various opportunities for instant learning feedback. For example, if the student conducts an activity that violates the process model, he could receive a message on a monitor next to the operating table. Technically, Celonis is already capable of both updating the imported data in real-time (pushing a new data model with every new activity tracked in the event log via live reloading) and sending out messages as well as recommendations in the case of certain process violations/triggers via the Action Engine [34]. This approach enables a live training feedback instead of the ex-post analysis presented in this report, which provides the opportunity for the student to improve errors and to act immediately during execution of the surgery training.

Violation weighting. The above-presented analysis can be extended by instructor-specific parameters. For example, instructors might put a larger weight on quality (better outcome) or time of the process (to be more cost-efficient and to conduct more surgeries per day). If this information is available, it can be used as input in a configuration sheet to be filled out first, and then the analysis adjust automatically according to these personal settings. Another valuable information would be a rating/flagging of violations. For example, the instructor could assign a number between 1 (violation is not so important) and 5 (violation is very important). These ratings could also be allocated to different areas of violations, like (i) safety, (ii) diagnosis, or (iii) surgery procedure itself. Based on this data, we could calculate weighted violation scores for each student.

7 Conclusion

Providing students and instructors with feedback on medical training performance is critical for the success of lab-based training programs. In this report, we show how process mining enables the analysis of event log data generated via video tracking of CVC installation trainings. We make use of the powerful analysis capabilities of process mining to generate meaningful and easy-to-use dashboards, which provide both students and instructors with fast and actionable insights.

Thereby, several objectives are considered: (i) Fast and clear overview of the actual training performance, (ii) intensive benchmarking between the training performance and the prescribed procedure for the CVC installation, and (iii) focus on relevant KPIs for the training group that are critical to derive training improvements. We also consider different levels of aggregation (stages and activities), as well as comparisons between the performed round (PRE and POST). This shows how process mining enables an end-to-end understanding of the CVC training procedure and data-based feedback.

The generated insights can be used to further innovate the data collection, the training itself, and the interpretation of deviations and errors in the process. This bears a lot of potential to achieve procedural training excellence, which is a key factor to oppose the critical shortage of staff and time in the healthcare sector.

References

1. Cooper, R.A. (2007): It's Time to Address the Problem of Physician Shortages: Graduate Medical Education is the Key. Annals of Surgery. 246(4), 527-534.
2. Fodeman, J., Factor, P. (2015): Solutions to the Primary Care Physician Shortage. The American Journal of Medicine. 128(8), 800-801.
3. Hawkins, M. (2017): Physician Supply Considerations: The Emerging Shortage of Medical Specialists, Merritt Hawkins White Paper Series.
4. World Health Organization (2006): Working together for Health: The World Health Report 2006.
5. Humphrey, H.J. (2010): Resources for Medical Education: Finding the Right Prescription. Transactions of the American Clinical and Climatological Association. 121, 76-93.
6. Herrmann-Werner, A. et al. (2013): "Best Practice" Skills Lab Training vs. a "see one, do one" Approach in Undergraduate Medical Education: An RCT on Students' Long-Term Ability to Perform Procedural Clinical Skills. PLoS ONE. 8(9), 1-13.
7. Kotsis, S.V., Chung, K.C. (2013): Application of the "See One, Do One, Teach One" Concept in Surgical Training. Plastic and Reconstructive Surgery. 131(5), 1194-1201.
8. Rohrich, R.J. (2006): "See One, Do One, Teach One": An Old Adage with a New Twist. Plastic and Reconstructive Surgery. 118(1), 257-258.
9. Qayumi, K. (2010): Surgical Skills Lab: A Hub for Competency Training. Journal of Investigative Surgery. 23(1), 48-56.
10. Vozenilek, J. (2004): See One, Do One, Teach One: Advanced Technology in Medical Education. Academic Emergency Medicine. 11(11), 1149-1154.

11. Smith, C.C. et al. (2004): Creation of an innovative inpatient medical procedure service and a method to evaluate house staff competency. Journal of General Internal Medicine. 19(5), 510-513.
12. Mason, W., Strike, P. (2003): See one, do one, teach one—Is this still how it works? A comparison of the medical and nursing professions in the teaching of practical procedures. Medical Teacher. 25(6), 664-666.
13. Ziv, A. et al. (2005): Simulation Based Medical Education: an opportunity to learn from errors. Medical Teacher. 27(3), 193-199.
14. Barrows, H.S. (1993): An overview of the uses of standardized patients for teaching and evaluating clinical skills. Academic Medicine. 68(6), 443-451.
15. Bradley, P., Postlethwaite, K. (2003): Setting up a clinical skills learning facility. Medical Education. 37(S1), 6-13.
16. Nikendei, C. (2005): Role-playing for more realistic technical skills training. Medical Teacher. 27(2), 122-126.
17. Issenberg, S.B. et al. (2005): Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Medical Teacher. 27(1), 10-28.
18. Jiang, G. et al. (2011): Learning curves and long-term outcome of simulation-based thoracentesis training for medical students. BMC Medical Education. 11(39), 1-7.
19. Lund, F. et al. (2012): Effectiveness of IV Cannulation Skills Laboratory Training and Its Transfer into Clinical Practice: A Randomized, Controlled Trial. PLoS ONE. 7(3), 1-10.
20. Ralhan, S. et al. (2012): Effective teaching skills—How to become a better medical educator. BMJ. 344, e765.
21. Strand, I. et al. (2009): Students Learning in a Skills Laboratory. Nordic Journal of Nursing Research. 29(3), 18-22.
22. Moran, J. et al. (2018): Current Technology in Advancing Medical Education: Perspectives for Learning and Providing Care. Academic Psychiatry. 42(6), 796-799.
23. Umer, R. et al. (2017): On predicting academic performance with process mining in learning analytics. Journal of Research in Innovative Teaching & Learning. 10(2), 160-176.
24. Mukala, P. et al. (2015): Learning Analytics on Coursera Event Data: A Process Mining Approach. Symposium on Data-Driven Process Discovery and Analysis (SIMPDA), Vienna, Austria.
25. Lira, R. et al. (2019): Tailored Process Feedback Through Process Mining for Surgical Procedures in Medical Training: The Central Venous Catheter Case. In: Daniel F., Sheng Q., Motahari H. (eds). Business Process Management Workshops. BPM 2018. Lecture Notes in Business Information Processing, Vol. 342. Springer, Cham.
26. Van der Aalst, Wil M.P. (2014): Process Mining Discovery, Conformance and Enhancement of Business Processes. Springer Berlin.
27. Munoz-Gama, J., de la Fuente, R., Sepúlveda, M., Fuentes, R., Conformance Checking Challenge 2019. 4TU.Centre for Research Data.
28. Corvetto, M. et al. (2017): Simulation-based training program with deliberate practice for ultrasound-guided jugular central venous catheter placement. Acta Anaesthesiologica Scandinavica 61(9), 1184–1191.
29. Carmona, J. et al. (2018): Conformance Checking. Relating Processes and Models. Springer Berlin.
30. Kerremans, M. (2018): Market Guide for Process Mining. Gartner. https://www.gartner.com/en/documents/3870291.
31. Veit, F. et al. (2017): The Proactive Insights Engine: Process Mining meets Machine Learning and Artificial Intelligence. In: Proceedings of the BPM Demo Track and BPM

Dissertation Award co-located with 15th International Conference on Business Process Modeling, Barcelona, Spain.
32. Yeole, A.S., Kalbande, D.R. (2016): Use of Internet of Things (IoT) in Healthcare. Proceedings of the ACM Symposium on Women in Research, Indore, India.
33. Geyer-Klingeberg, J. et al. (2018): Process mining and robotic process automation: A perfect match. In: Proceedings of the Dissertation Award, Demonstration, and Industrial Track at BPM 2018 co-located with 16th International Conference on Business Process Management, Sydney, Australia.
34. Badakhshan, P. et al. (2019): The Action Engine - Turning Process Insights into Action. In: Proceedings of the Business Process Management (BPM) Demo Track and BPM Dissertation Award, Vienna, Austria.