DTU Library

Analysis of Emergency Room Episodes Duration Through Process Mining

Rojas, Eric; Cifuentes, Andres; Burattin, Andrea; Munoz-Gama, Jorge; Sepúlveda, Marcos; Capurro, Daniel

Published in:
Business Process Management Workshops - BPM 2018 International Workshops, Revised Papers

Link to article, DOI:
10.1007/978-3-030-11641-5_20

Publication date:
2019

Document Version
Peer reviewed version

Citation (APA):
Rojas, E., Cifuentes, A., Burattin, A., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2019). Analysis of Emergency Room Episodes Duration Through Process Mining. In F. Daniel, Q. Z. Sheng, & H. Motahari (Eds.), *Business Process Management Workshops - BPM 2018 International Workshops, Revised Papers* (pp. 251-263). Springer Verlag. https://doi.org/10.1007/978-3-030-11641-5_20

Analysis of Emergency Room Episodes Duration through Process Mining

Eric Rojas[1,2], Andres Cifuentes[1], Andrea Burattin[3], Jorge Munoz-Gama[1], Marcos Sepúlveda[1], and Daniel Capurro[2]

[1] Department of Computer Science. School of Engineering
[2] Department of Internal Medicine. School of Medicine
Pontificia Universidad Católica de Chile
{eric.rojas,alcifuen,jmun}@uc.cl
marcos@ing.puc.cl - dcapurro@med.puc.cl
[3] Technical University of Denmark, Kgs. Lyngby, Denmark

Abstract. This study presents the proposal of a performance analysis method for ER Processes through Process Mining. This method helps to determine which activities, sub-processes, interactions and characteristics of episodes explain why the process has long episode duration, besides providing decision makers with additional information that will help to decrease waiting times, reduce patient congestion and increment quality of provided care. By applying the exposed method to a case study, it was discovered that when a loop is formed between the Examination and Treatment sub-processes, the episode duration lengthens. Moreover, the relationship between case severity and the number of repetitions of the Examination-Treatment loop was also studied. As the case severity increases, the number of repetitions increases as well.

Keywords: Process mining · Healthcare · Emergency Room

1 Introduction

Performance measurements in Emergency Room (ER) processes are highly important, because of the information they can provide to identify behaviour of episodes with extended waiting times for patients expecting attention. Generally, patients are categorized into triage categories, normally by a nurse, to determine their attention waiting times in the ER. In this ER, the Manchester triage [1] is used to classify patients in five color categories: red, orange, yellow, green and blue. Red being the most critical patients with lowest waiting times, while blue being the least severe patients with highest waiting times. Identifying improvement opportunities in the ER processes can help reduce waiting times, improve the quality of provided services and reduce overcrowding [2].

Our approach to study the performance of ER processes is based on process mining [3] as the main component to identify, discover and analyze activities executed during ER process episodes. Process mining is a research discipline that focuses on extracting process knowledge from data generated and stored in the databases of (corporate) information systems, in this case Hospital Information Systems (HIS) [3,4,5].

Process execution data is extracted as event logs, which are sets of episodes, each containing all the activities executed for a particular process instance. Process mining tools and techniques can be applied to discover process models, verify conformance, analyze organizational patterns and check the performance of a process in any hospital [3]. This paper proposes a process mining methodology to study the performance of ER processes. These processes have been analyzed using simulation, data mining along with process mining [6,7], among others. Data mining has been used to find patterns and understanding the causes of certain process behavior while in the other hand process mining describes how processes are currently performed.

Previous studies describe how process mining has been applied to analyze the executed processes within ER [7,8,9]. These studies have given insight about the process and the flow of activities during episodes (for example, medication or discharge activities). The first case is a study in a Portugal hospital where a software suite was defined to extract data, build an event log and discover any process in the medical center [7]. An specific case study was done using ER data, but the solution is general. The solution includes clustering techniques and Markov chain models. The second study is exploratory and was conducted in four Australian hospitals [8], where data was extracted, an event log was built, and discovery techniques of process mining were applied. The third one proposes a methodology based on frequently posed questions, and provides a case study, but no performance analysis was executed [9]. Performance analysis of ERs has been previously researched [10,11], but did not included a method for performance analysis of the ER episodes or any ER metric in general using process mining techniques.

The objectives of this paper are to analyze the ER episodes behavior using a process mining methodology comprehensively described, to apply it to a case study, and to determine which activities, sub-processes and their interactions in the ER process explain why the process get stalled and has a longer duration, and, to identify any existing relationships between some characteristics of the process activities or sub-processes and the process performance. This paper is an exploratory study of performance analysis in ER using process mining. Further studies must be done in order to complete more in depth quantitative analysis.

The structure of the paper is as follows: Section 2 describes the proposed methodology. Section 3 describes a case study where the methodology has been applied including results and discussion. Finally, conclusions and future work are highlighted.

2 Method

A 6 phases methodology for the process performance analysis of ER episodes is proposed, as depicted in Figure 1 [1]. To generate this methodology previous studies from the authors were considered [9].

Phase 1: Extraction and Transformation. ER processes are supported by clinical and non-clinical activities that are executed by different types of resources (physicians, nurses, administrative staff). Each of these series of activities correspond to an episode,

[1] All figures presented in this paper can be seen with more details in the following link `https://wp.me/p9ZIAl-1g`

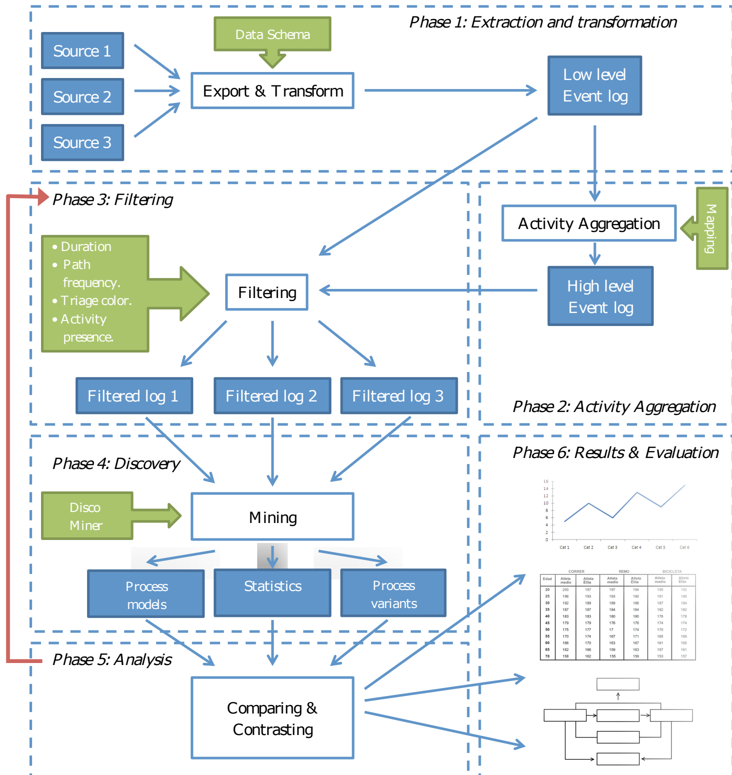

Fig. 1: The 6 phases methodology adopted in this paper.

which is registered in any Hospital Information System (HIS) [4]. HIS are computer systems designed to ease management of the hospital's medical and administrative information and to improve quality of healthcare [12]. HIS store records as events (or activities) that include all the necessary data to create an event log to perform process mining analysis. An event log is a file record that provides an audit trail that can be used to understand the system activities. In this phase, HIS records are extracted from different sources (databases, repositories, etc.) and transformed into a standard event log format that is readable by process mining tools like CSV, XLS, XES, among others.

Phase 2: Activity Aggregation. The complexity of the ER process is defined by the amount of different activities involved in one process instance and the lack of process structure: a classic spaghetti process [3]. Spaghetti processes are unstructured models with several activities connected to each other and are difficult to visualize because of the amount of information that is shown when represented as a graphical model. Thus, it becomes necessary to bring the log to a higher abstraction level, with a general view of the process in terms of flow and structure, by shifting from a perspective which is data-oriented to a process-centric one. The result is an event log that contains sub-processes rather than activities as its basic unit, which makes inspection and analysis more accessible. The followed approach aggregates activities by mapping them to the sub-process they belong to. Sub-processes are defined by some activities that work as separators and, in case the separation is not evident, process semantics. This mapping can be done by manually correlating each activity to its corresponding sub-process. Afterwards, an activity name substitution following the mapping will have a high-level event log as a result. The activities of this new event log will be the sub-processes comprising the activities of the old low-level log. Reducing the amount of different activities results into a simpler process model.

Phase 3: Filtering. After aggregating activities into sub-processes and generating the high-level log, the next step is filtering the log to reduce noise. There are several activities that are either non relevant to the entire process (but are still registered because they contain descriptive information) or are added to the historical record only for information storage purposes, thus not giving additional process input. These activities are identified by using expert knowledge to directly discard them and by filtering out activities that follow highly uncommon paths or have a low frequency. As it is shown in Figure 1, Phase 3 is revisited later to generate specific process event logs depending on the goal of the process analysis. The filtered high-level log retains all attributes coming from the low-level source log. Using the attributes enumerated at the event log schema it is possible to filter the event log to generate sub-logs that lead to specific process models that can answer questions related to process behaviour depending on Triage color, Diagnosis, length of the stay (duration), frequency of a path, among others.

Phase 4: Discovery. Using the filtered event logs, process models are generated in Disco [1] (or any other process mining tool, e.g. ProM[2]) to be able to look in depth and analyze the behaviour of the process. Depending on the filters applied at Phase 3, different models are obtained. In this phase, numerical information can be collected for different process variants. The goal is to collect data of the process variants to elaborate a comprehensive analysis during the next phase. To build process models event logs, process mining discovery algorithms go through the log and maps it onto a process model such that the model is "representative" of the behavior seen in the log [3].

Phase 5: Analysis. Inspecting the process model makes it easier to identify sub-processes by finding exit activities (i.e., activities that are frequently the end of most paths of a section of a process). Before analyzing it is necessary to get the correct model by filtering the log according to what is going to be analyzed and then discovering a new process model, which means that there is a need to go back to Phase 3 and 4 to filter the log and generate the corresponding model. Figure 1 shows this procedure with an arrow starting at Phase 5 and going back to Phase 3, representing the possibility to iterate through Phases 3, 4 and 5 to perform a comprehensive analysis of the process behavior.

Phase 6: Results and Evaluation. This last step considers the delivery of process models, summarizing the data extracted from process models and comparison charts for

[1] See `http://www.fluxicon.com/disco`
[2] See `http://www.promtools.org`.

process model duration, and performing validation and evaluation of these results and conclusions through expert opinion.

3 Case Study

Using the methodology described in the previous section, a case study was conducted at the Clinical Hospital of Red de Salud UC CHRISTUS, using historical data collected at its ER during July 2014. The hospital is located in Santiago, Chile. Figure 2 shows a BPMN diagram[3] of the expected high-level flow of the process according to experts. A Triage sub-process and a Treatment sub-process were identified. Once a Triage sub-process was found, a deeper analysis using the triage color attribute was performed. After the Triage sub-process finishes, the rest of the process corresponds to Treatment and Discharge sub-processes. Identifying the Discharge sub-process is simple since discharging a patient is a straightforward action. The Treatment sub-process involves most of the activities for each case. The conversion from low-level to high-level log will collapse most activities depending on the nature of the tasks performed by the resources of the ER. There are examination activities that figure out what is the problem of the patient based on physical examination and tests, and to confirm if the patient is ready to be discharged. There are also treatment activities that are performed to lead the patient to a stable condition. These collapsed activities represent the Examination for Prediction, Treatment and Examination for Validation sub-processes. Whenever the treatment is not successful, another instance of Examination for Prediction starts, followed by Treatment. In some cases, Examination for Validation is performed by medical order. Analyzing the duration of these sub-processes provides insight about the behaviour of the process and how performance of these sub-processes affect the whole process. Analyzing the Treatment models through filtering by duration defines differences between short stays and long stays. Finally, considering referrals to specialized physicians provides information about duration of the process depending on the complexity of the diagnosis. A detailed description of the executed tasks performed at each Phase of the study is provided next.

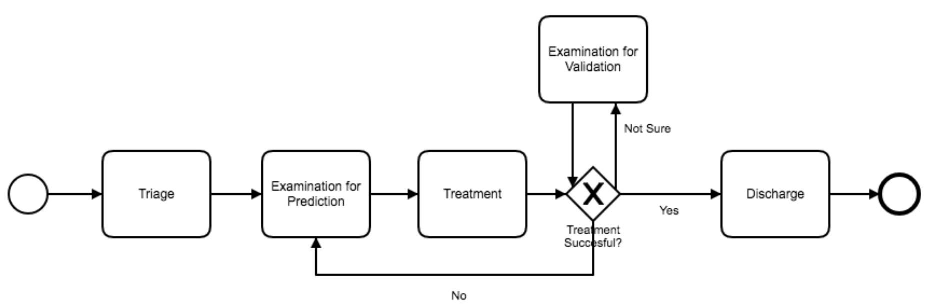

Fig. 2: ER process: high-level BPMN model with the expected high-level process flow.

[3] See `http://www.bpmn.org`.

Phase 1: Extraction and Transformation. The data used to construct the event log is historical data collected by the research team from the Hospital Information Systems Alert ADW Phase I, which corresponds to the system used to store the data in the ER Process. Historical data was built as an enumeration of 309,796 activities registered, composing 7,160 different episodes. There are 64 different activity types registered. Each register includes its timestamp, its resource and the episode id. In addition to these attributes, the analysis considers diagnosis, type of resource (nurse, doctor, technician, auxiliary nurse) and triage color. It is worth noting that the activity timestamps are registered at the hour level. However, timestamps of the first and the last activity are registered considering minutes and seconds. By selecting corresponding attributes using Disco import tool, an event log is built based on this historical data.

Phase 2: Activity Aggregation. Disco was used to filter the event log according to medical expert opinion, who explained that the activities could be classified in 3 sub-processes: Triage, Examination-Treatment and Discharge activities. These groups are initially considered as sub-processes of the ER process. Each of the original activities is mapped to a collapsed activity by enriching the event log with this information. Table 1 shows the relationships created to achieve this representation for the Treatment sub-process. Activities that are not included in this table are disregarded because they are not providing relevant information to the process model, according to the expert.

Concerning the Examination sub-process, activities could be of two types: "for prediction" or "for validation". Examination for prediction is the first examination, performed by a doctor, that provides enough information *to figure out* what problem the patient has. Examination for validation is an activity, also performed by a doctor, that is done with the goal *to check* that everything is fine with the patient and that there is no need to continue staying at the ER. To make a difference between both types it is necessary to check the whole current episode. If current examination is the last one – in the specific episode –, and it is not the first one, it is considered an examination for validation, else it is an examination for prediction. Identifying types of examination was performed using a script implemented for this purpose. The described procedure is formalized in Algorithm 1.

Phase 3: Filtering. At the first filtering phase, events that were not mapped to any sub-processes are removed using Disco filtering capabilities. This filter removes noise and paths that are not required to analyze due to their low frequency.

Table 1: Mapping of activities into sub-processes for the Treatment process.

Original Activities	**Sub-process**
Nurse task and Physician task	Examination
Prescribe Medication, Perform Procedures, Physical Examination, Give Medication, Prescribed Procedures, Prescribed Medication (internal), Required Laboratory Tests, Required Imagenology Tests, Biometry, Other Required Tests, and Cancelled External Medication Prescription	Treatment
Clinic Discharge and Last Discharge	Discharge

Algorithm 1: Identify examination type

```
begin ExaminationType(log)
    forall episode ∈ log do
        for i = 0 up to |episode| do
            if episode[i].name is "Examination" then
                lastExamination ← TRUE
                for j = i + 1 up to |episode| do          // |episode| is ep. length
                    if episode[j].name is "Treatment" then
                        lastExamination ← FALSE
                        break
                if lastExamination then
                    set examination type for episode[i] to "validation"
                else
                    set examination type for episode[i] to "prediction"
```

Phase 4: Discovery. Once the filter is applied, Disco generates a process model that represents the portion of the log resulting from the event log aggregation and filtering [13]. Disco also provides performance analysis, but this did not provide significant results for the analysis. From the aggregated event log, we obtain the process model illustrated in Figure 3. The generated model is similar to the one previously shown in Figure 2. The main difference between the models is that Figure 3 includes infrequent paths (e.g., episodes where the patient is taken back from Discharge to Treatment).

The Triage sub-process normally starts with a "Nurse task" activity. After that, the process continues to a "Physician task" and then to an "Intake task" activity. In a third of the cases, a doctor does not take part at the triage sub-process and it goes directly to "Intake task". About 10% of the cases that go through the doctor are then taken by a technician ("Technician task"), and then an intake note is generated. After the generation of the intake note, a physical examination could occur or vital signs could be registered (these activities could be skipped). Finally the first triage is done, and the triage sub-process is finished.

The Treatment sub-process begins with "Examination for Prediction". This activity includes all nurse and doctor activities that occur before any actual treatment is performed. After this examination, a consultation to a specialized doctor could happen. Then, the process continues with the actual treatment. After the treatment is finished, the patient is discharged. Also, a loop between examination for prediction and treatment is identified. And, finally, after the patient is treated, a doctor could request a final examination to validate whether the patient could be discharged or not. It is worth noting that the difference between short and long episodes depends almost exclusively of this sub-process. Long episodes of the complete process are the ones that include several of these "examination for prediction - treatment" loops.

Finally, patient discharge is straightforward. After aggregating the activities into one general discharge activity, the sub-process was collapsed into only one activity. The process structure is simple and sequential. Each activity is directly followed by the previous one with no alternative paths.

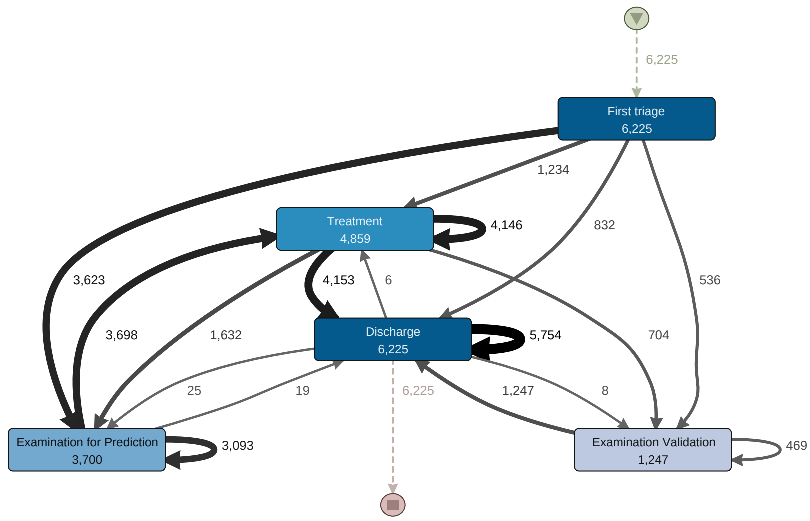

Fig. 3: Process model for the enriched event log. Arrows are decorated with the episode frequency of each path.

Phase 5: Analysis. The analysis of the Triage sub-process was done separately since it is the starting point of the ER process, and the examination, treatment and discharge sub-processes were analyzed together to identify loops and possible relationships between them, given that these three sub-processes happen at the attention box and are performed by the same resources.

Triage sub-process. The initial sub-process in each ER episode consists in determining the severity of each episode according to the patient conditions. Usually, it consists in categorizing each patient to a Manchester triage category(red, orange, yellow, green and blue) [1]. Several activities were selected by the ER specialist from the event log to see the process model followed in the analyzed episodes. Figure 4 shows a process model with the activities executed during the Triage sub-process. For a more detailed analysis of the process, the event log was split in two groups: one half with the fastest episodes and the other half with the slowest episodes. No relevant differences in activities presence or their sequence were discovered.

Examination, Treatment and Discharge Sub-process (ETD). The second sub-process includes three main subsets of activities. The first subset contains the activities related to the examination of the patient. The second subset is the treatment, including treatment activities, medications and exams. Finally, discharge activities, including paperwork and the actual patient discharge, are grouped into a third subset. Figure 5 shows a process model with the activities executed during the ETD sub-process.

For a more detailed analysis of the process, the event log was split into two groups. The first group corresponds to the half with the fastest episodes. In Figure 6 the pro-

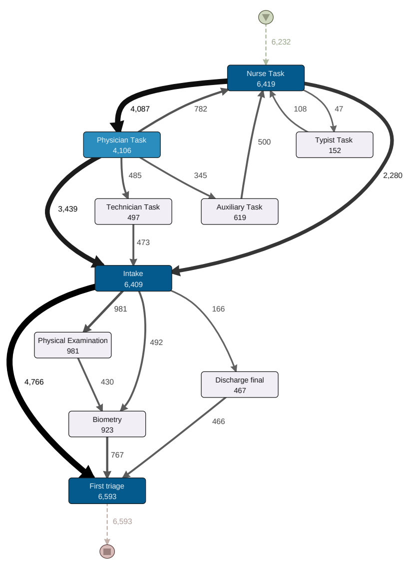

Fig. 4: Triage Sub-process Model

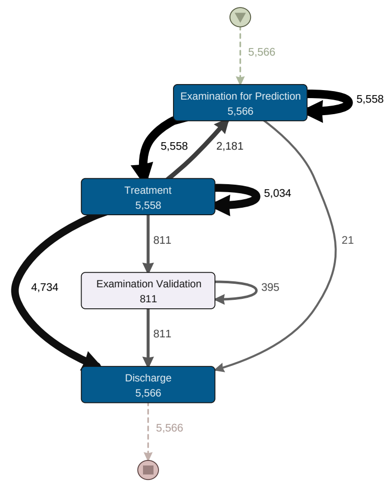

Fig. 5: ETD sub-process

cess model for these episodes is shown. The other group includes the half with the slowest episodes, and its process model can be seen in Figure 7. A loop between examination for prediction and treatment was identified on both Figure 6 and Figure 7. This loop could occur several times in one instance of the process. Every time this loop is followed, the process time increases (considering a discrete average of one hour), impacting on the episode duration. 20% of the cases have a examination for validation before discharging a patient. By inspecting the average duration (discrete average of one hour at treatment-examination validation transition, and another hour for examination validation-discharge transition) and considering that activity times are registered by the hour, it is safe to assume that examination for validation extends from one to two hours, and most of the time it is found at long process instances. Although both groups have the same activity structure, there is a great duration difference between both of them. The reason is that most of fastest cases don't loop back to examination for prediction, while most of slowest cases do.

Analysis by triage color. A relevant case attribute for this process is triage color since it informs resources about the severity of the patient case [1,14]. Table 2 shows a detailed description of the characteristics of the loop mentioned in the previous analysis. Higher priority episodes have a higher average of repetitions (red and blue episodes are not considered because they correspond to the least amount in the sample). Orange episodes involve more complex diagnosis and treatments which explains that in average these cases have more than two repetitions of the loop. While the severity of the episode becomes higher, the average number of repetitions increases as well. Green episodes, in average, have half the number of repetitions (1.27 average repetitions) of

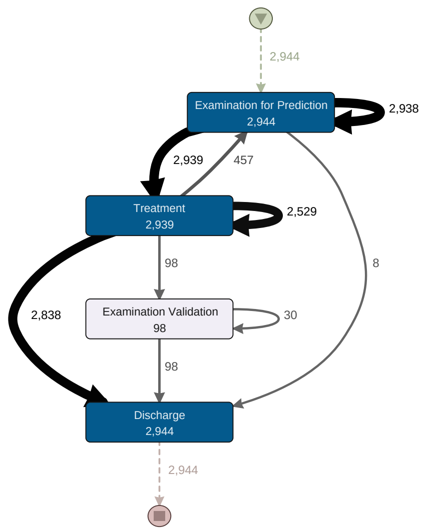

Fig. 6: ETD - fastest cases

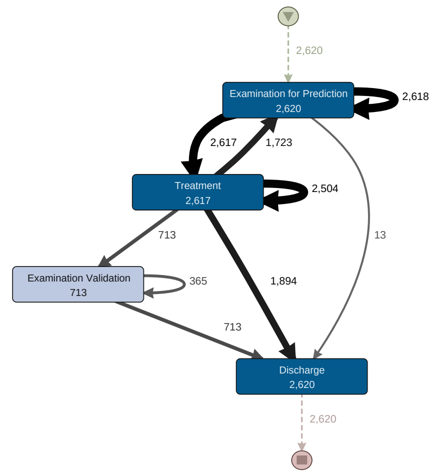

Fig. 7: ETD - slowest cases

orange episodes (2.34 average repetitions). Figure 8 show these results. The chart has a value mark at the maximum number of repetitions. The chart is also divided in quartiles. Figure 8 includes all episodes of the event log and the number of repetitions of each episode plotted in black. By simply inspecting the chart, it is clear that more than one third of total episodes have at least two repetitions of the Examination-Treatment loop. When analyzing by color the proportion of episodes with two or more repetitions, this proportion increases as the severity of the case gets higher. This trend is shown in Figure 9. The trend of higher severity cases is more steep, which reinforces previously shown results. For example, the trend for green cases moves between one and two repetitions, while the trend for orange cases is between one and six repetitions.

Phase 6: Results and Discussion. By analyzing the generated process models with ER experts, there are two places where the ER process slows its pace, increasing waiting times. The first situation is at the examination-treatment sub-process. Depending on how many times the patient is examined to detect what his/her problem is, the process extends per each examination-treatment loop, impacting total duration. The other section where the process takes longer than expected is when examination for validation takes place, extending the duration by one up to two two hours (on average). There is a relationship between triage color and the number of examination-treatment loops. As the severity of the case increases, the number of repetitions increases as well, showing that blue and green cases (low severity) have a below average number of repetitions while yellow, orange and red cases (high severity) have a number of repetitions above the average. Triage color does not affect any other sub-process of the ER process.

Table 2: Examination for prediction - treatment loop data by triage color

	All	Blue	Green	Yellow	Orange	Red
Episode frequency	5538	44	2347	2374	755	18
Absolute frequency	9184	51	2292	4328	1773	40
Relative frequency	100%	0.56%	32.58%	47.1%	19.3%	0.44%
Max repetitions	13	3	9	10	13	7
Average repetitions	1.66	1.16	1.27	1.82	2.34	2.22
1 or more repetitions	99%	100%	99%	99%	99%	100%
2 or more repetitions	39%	13%	21%	29%	64%	61%
3 or more repetitions	14%	2%	4%	19%	32%	27%

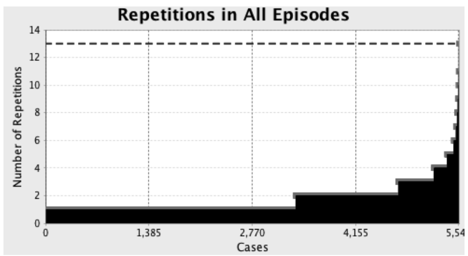

Fig. 8: All Episodes

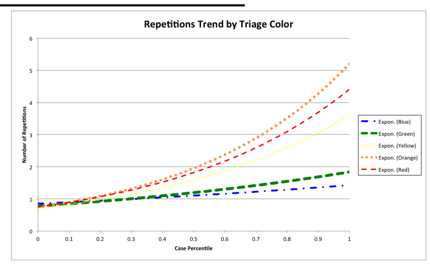

Fig. 9: Episode Repetitions Trends by Triage Color

Discussion In this paper, several sub-processes and their relationships were identified. From the ER point of view of the expert, this helps to see which activities implicate an increment on the episode duration. The increment of episode duration directly increases the waiting times for all the episodes, and increments the ER overcrowding, which has become one of the most relevant issues nowadays [15].

Four sub-processes were identified, but two main sub-processes are critical. First, the Examination sub-process, that includes activities where the resources try to determine the diagnostic of the patient to take actions to bring care to the patients. And secondly, the Treatment sub-process where any exam, procedure or medication is provided. In this research, through process mining, a loop between the two sub-processes was identified. Every time it happens, time is added to the episode duration. It was also analyzed according to the different triage colors, giving a more detailed description on when this loop happens more often, so as to pay more attention in future episodes.

From the process mining perspective, the methodology was significant to the field. Two levels of analysis were executed, low level detailed analysis was performed with the domain expert to identify the executed activities and the sub-processes to which they belong, and high level analysis was carried out to see the relationships between the different sub-processes to identify which are the causes of slowest episodes.

4 Conclusion and future work

By analyzing the ER process, it was discovered that the loop between Treatment and Examination sub-processes increments the duration of ER episodes. This will help with the identification of the episodes where this happens (mainly by triage color), to make any necessary improvements. Identifying these loop as the main cause of the increment in the episode duration time is significant to help reduce it and with this reduction, lower the episode times, free boxes in the ER, give a faster attention to more patients, shorten waiting times and finally reduce the overcrowding of the ER. Further work will include applying additional process mining and more statistical techniques to analyze the ER process and complete more in depth quantitative analysis. Besides more advanced research of the evidences provided by these models should be conducted in the ER.

References

1. Manchester Triage Group. *Emergency Triage*. BMJ Publishing, 1997. 1, 8, 9
2. R Konrad, K DeSotto, A Grocela, P McAuley, J Wang, J Lyons, and M Bruin. Modeling the impact of changing patient flow processes in an emergency department: Insights from a computer simulation study. *Operations Research for Health Care*, 2(4):66 – 74, 2013. 1
3. Wil M. P. Van der Aalst. *Process Mining: Data Science in Action*. Springer, 2016. 1, 2, 3, 4
4. Ronny Mans, Wil M. P. Van der Aalst, and Rob JB Vanwersch. *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*. Springer, 2015. 1, 3
5. Eric Rojas, Jorge Munoz-Gama, Marcos Sepúlveda, and Daniel Capurro. Process mining in healthcare: A literature review. *Journal of Biomedical Informatics*, 61:224 – 236, 2016. 1
6. Alexander Kolker. Process modeling of emergency department patient flow: Effect of patient length of stay on ed diversion. *Journal of Medical Systems*, 32(5):389–401, 2008. 2
7. Álvaro Rebuge and Diogo R. Ferreira. Business process analysis in healthcare environments: A methodology based on process mining. *Information Systems*, 37(2):99–116, 2012. 2
8. Andrew Partington, Moe Wynn, Suriadi Suriadi, Chun Ouyang, and Jonathan Karnon. Process mining for clinical processes: A comparative analysis of four australian hospitals. *ACM Transactions on Management Information Systems (TMIS)*, 5(4):1–19, 2015. 2
9. Eric Rojas, Marcos Sepúlveda, Jorge Munoz-Gama, Daniel Capurro, Vicente Traver, and Carlos Fernandez-Llatas. Question-driven methodology for analyzing emergency room processes using process mining. *Applied Sciences*, 7(3):302, 2017. 2
10. Shari J Welch, Brent R Asplin, Suzanne Stone-Griffith, Steven J Davidson, James Augustine, Jeremiah Schuur, and Emergency Department Benchmarking Alliance. Emergency department operational metrics, measures and definitions: results of the second performance measures and benchmarking summit. *Annals of emergency medicine*, 58(1):33–40, 2011. 2
11. Christian Michel Sørup, Peter Jacobsen, and Jakob Lundager Forberg. Evaluation of emergency department performance-a systematic review on recommended performance and quality-in-care measures. *Scand J Trauma Resusc Emerg Med*, 21(62):1–14, 2013. 2
12. P. Degoulet. *Hospital Information Systems*, pages 289–313. Springer Paris, Paris, 2014. 3
13. C. W. Günther and A. Rozinat. Disco: Discover your process. *BPM (Demos), 2012*, 2012. 7
14. Kevin Mackway-Jones and Colin Robertson. Emergency triage. *BMJ-British Medical Journal-International Edition*, 314(7086):1056, 1997. 9
15. Nathan Hoot and Dominik Aronsky. Systematic review of emergency department crowding: causes, effects, and solutions. *Annals of emergency medicine*, 52(2):126–136, 2008. 11