Journal of Biomedical Informatics 78 (2018) 60–77

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Discovering role interaction models in the Emergency Room using Process Mining

Camilo Alvarez[a], Eric Rojas[a,*], Michael Arias[a], Jorge Munoz-Gama[a], Marcos Sepúlveda[a], Valeria Herskovic[a], Daniel Capurro[b]

[a] Computer Science Department, School of Engineering, Pontificia Universidad Católica de Chile, Chile
[b] Internal Medicine Department, School of Medicine, Pontificia Universidad Católica de Chile, Chile

A R T I C L E I N F O

Keywords:
Healthcare
Processes
Process mining
Case studies
Organizational mining
Organizational team patterns

A B S T R A C T

Objectives: A coordinated collaboration among different healthcare professionals in Emergency Room (ER) processes is critical to promptly care for patients who arrive at the hospital in a delicate health condition, claiming for an immediate attention. The aims of this study are (i) to discover role interaction models in (ER) processes using process mining techniques; (ii) to understand how healthcare professionals are currently collaborating; and (iii) to provide useful knowledge that can help to improve ER processes.
Methods: A four step method based on process mining techniques is proposed. An ER process of a university hospital was considered as a case study, using 7160 episodes that contains specific ER episode attributes.
Results: Insights about how healthcare professionals collaborate in the ER was discovered, including the identification of a prevalent role interaction model along the major triage categories and specific role interaction models for different diagnoses. Also, common and exceptional professional interaction models were discovered at the role level.
Conclusions: This study allows the discovery of role interaction models through the use of real-life clinical data and process mining techniques. Results show a useful way of providing relevant insights about how healthcare professionals collaborate, uncovering opportunities for process improvement.

1. Introduction

The aim of the Emergency Room (ER) is to assess and treat patients, and provide medical care so they can recover or, at least, alleviate any presented illness or set of symptoms. This can be achieved through standard and well executed processes in the ER, where healthcare professionals collaborate in a systematic manner. Healthcare professionals include physicians, nurses, technicians and other personnel who work together to make quick and accurate decisions for the recovery of the patient. Team interactions (collaborative work) have become a key aspect for successful ER results.

Studies have shown that inadequate team interaction in ER has been a primary or contributing cause of more than half of malpractice claims [1,2]. Inadequate team structure, poor communication patterns, inappropriate setting of goals and responsibilities for each role, lack of standard protocols, poor prioritization of activities, inappropriate management of clinical knowledge and skills, bad coordination of care across patient conditions, services, and settings over time and non-involvement of the relevant actors in the decision making process can directly impact the results of patients' treatment in the ER [1,3]. On the other hand, several advantages can be obtained when a good team interaction is kept in consideration when planning and executing ER processes. The most significant one is that adequate team interactions can reduce patient morbidity and mortality. Moreover,this can save economic resources, reduce clinical error rates, reduce legal costs, and even impact other human aspects such as reduce stress an frustration among healthcare professionals [1,2,4,5].

Accordingly, this research seeks to understand and study the limited part of the communication and interaction that is carried out via documentation in the Hospital Information Systems (HIS). This study does not include in its scope other interactions such as verbal and non-verbal communications. To achieve this understanding, the interaction among resources must be studied. With the arrival of Process Mining and Data mining to the healthcare domain, not only how processes are performed can be studied in detail, but also how the different individuals and roles interact.

* Corresponding author.
E-mail addresses: cealvarez@uc.cl (C. Alvarez), eric.rojas@uc.cl (E. Rojas), m.arias@uc.cl (M. Arias), jmun@ing.puc.cl (J. Munoz-Gama), marcos@ing.puc.cl (M. Sepúlveda), vherskov@ing.puc.cl (V. Herskovic), dcapurro@med.puc.cl (D. Capurro).

https://doi.org/10.1016/j.jbi.2017.12.015
Received 29 June 2017; Received in revised form 23 November 2017; Accepted 28 December 2017
Available online 28 December 2017
1532-0464/

Process mining is quicker than observation studies because it uses data already available and algorithms that can quickly analyze large datasets. Process Mining can analyze data about all interactions recorded in (HIS). Therefore, it is more comprehensive than observation since it can use data about all episodes recorded during an extensive period of time. On the other hand, observation can handle a broader span of data, since it is not limited to the interactions recorded in the HIS. Both methods can supplement each other to gain a broader understanding of the interactions. Through triangulation research methods, the results obtained using both process mining and observation could be checked and validated [6].

Therefore, the research question of this article is *how to discover role interaction models in ER through process mining techniques?*

To find these role interaction models, a four step method is proposed: first, centered on the extraction of the required data; second, the retrieved data are handled in order to build an event log that stores the organizational data for further analysis; third, process mining techniques are applied to discover role interaction models; and fourth, the results are validated with ER experts. Conducting a case study in an high-complexity emergency room of one of the main university hospitals in Chile, where organizational interaction models were discovered, was used as a way to validate the method.

This article contributes to this research field by introducing a novel method that use process mining tools and techniques to discover role interaction models, the use of filters based on ER episode attributes to analyze them, and the validation of the proposed method in a real case study. The use of the proposed method provides ER healthcare professionals with a set of tools to analyze more ER episodes in a shorter period, provide objective and replicable analysis based on executed episodes, different levels of analysis in an easier and faster way, and the use of filters to analyze subsets of episodes that represent different scenarios in ER. The application of the proposed method in a real case study has generated relevant insights about how healthcare professionals collaborate.

This article is structured as follows. Section 2 describes the background of this research, including Process Mining in Healthcare and the Organizational Perspective. Section 3 presents the proposed method for role interaction model discovery. Section 4 illustrates the application of the proposed method through a case study. Section 5 describes the obtained results, and Section 6 includes the discussion of these results from a healthcare perspective. Finally, Section 7 presents conclusions and future work.

2. Related work

2.1. The emergency room

As defined by the American College of Emergency Physicians, "Emergency medicine is the medical specialty dedicated to the diagnosis and treatment of unforeseen illness or injury. …The practice of emergency medicine includes the initial evaluation, diagnosis, treatment, coordination of care among multiple providers, and disposition of any patient requiring expeditious medical, surgical, or psychiatric care."[1] The ER is the physical location where physician, nurses, technicians, administrative resources and healthcare professionals in general work towards assisting these patients.

The usual first contact in an ER is the nursing staff, which are trained to execute the screening examination, in order to classify and provide attention to the different patients, regarding their severity. After the patients have been initially screened, the healthcare professionals provide their assistance.

In literature and in practice several names are used to define the ER. These additional names include: emergency clinic [7], accident and emergency [8], emergency department, among others.

2.2. Process oriented data science

With the appearance of Hospital Information Systems (HIS) that support the execution of processes in hospitals, new data is available and new techniques can be used to analyze it. Terms such as big data, prediction, expert systems, data analytics are now a significant part of the tools that experts use to obtain a more exhaustive analysis of the available data [9,10]. All these new techniques can be grouped under the term Data Science [11], that for many is the profession of the future. Several of these techniques can be applied to analyze the activities performed in any ER: process mining, data mining, machine learning and artificial intelligence, among others. Process oriented data science can be used within the ER as well as within all other medical units. The main subject in this research is process mining, but additional data science fields could also be applied to complement the analysis.

2.3. Process mining

Within the healthcare domain, large amounts of data are recorded regarding the diagnosis and treatment of patients [12,13]. This information represents an opportunity to discover and validate if standards are followed This information represents an opportunity to discover and validate if standards are followed and understand how healthcare processes are performed [3,14]. Improving ER processes is difficult because of their well-known complexity [14], and the fact that each patient is a different episode and there are many variables, such as diagnosis, severity and medications, that can affect the appropriate treatment. Moreover, the collaboration among healthcare professionals might have a determinant impact in the evolution of the patient.

Process mining is a relatively young research discipline that focuses on extracting knowledge from data stored in the databases of (corporate) information systems in order to build event logs [11]. An event log can be viewed as a set of traces, each containing all the activities executed for a particular process instance or, in our case, an episode. An episode is a series of events that happen to a person and a series of activities a group of specialists undertake in order to treat that person [15].

Process-Aware Information Systems (PAIS) are systems that produce event logs [16]. Specific examples of such applications include Enterprise Resource Planning systems (e.g., SAP[2]), and Hospital Information Systems (HIS [17]). Event log data are not limited simply to the data from these applications, as many other systems can also provide useful information about the executed process.

There are three main types of process mining: process discovery, conformance checking, and enhancement. Automatic process discovery allows process models to be extracted from event logs, conformance checking allows monitoring deviations by comparing the event log with a given model, and enhancement allows extending or improving an existing process model [11]. Additionally, there are four different process mining perspectives [11]: the control-flow perspective considers the control-flow of activities, the organizational perspective focuses on information about resources and how they are related during process execution, the case perspective places special attention on the properties of cases; and the time perspective focuses on analyzing the process from a temporal point of view. Most process mining research focuses on the control-flow perspective [18–20]. According to [21], some research has also focused on performance analysis [22,23].

[1] https://www.acep.org/Clinical—Practice-Management/Definition-of-Emergency-Medicine.

[2] http://go.sap.com.

2.4. Organizational perspective

The organizational perspective aims to discover the relationships among resources, to explore if they collaborate among themselves, and to discover the existence of social networks among them. Within this perspective, it is possible to take into account the information about resources in order to enrich the process analysis [24], study team and resource allocation [25,26], and also to use this information to analyze organizational patterns [27]. For example, Russell et al. [28], have identified a number of Workflow Resource Patterns, which aim to represent the different ways in which resources are utilized in workflows.

These patterns capture information related to the roles, their capabilities, and their history, among other criteria. However, these patterns do not consider the collaboration between resources, a key aspect that needs to be studied further on a human task level [29], especially in scenarios that include highly collaborative processes, like ER. Previous work on task-based collaboration scenarios [29], elaborate a social network analysis in order to provide insights about the collaboration between teams of different departments in a Dutch hospital. This analysis focuses on determining the interaction among departments, and not among the people involved in the process. Human resources and the teams involved in ER processes are worthy of analysis, in order to discover role interaction models that could be use to improve patient care in the future.

Recently, some of the authors of this paper conducted a literature review on the application of process mining in healthcare [21], where more than 70 case studies have been identified, reviewed and categorized. It provides an overview of the state of the art and the benefits, importance, limitations and future trends on this field. However, to the best of our knowledge, there are no studies related to the discovery of role interaction models that have been performed in ER processes using process mining.

Several works in the literature have illustrated how process mining could complement existing observational techniques in the HIS analysis, or could be applied in situations where observation is not possible or desirable [21]. For example, process mining analysis can involve a long period of time, can be replicated periodically at low cost, and minimize the observation bias since data are extracted from the HIS [11]. However, although the academic scope of the paper only addresses the application of the proposed method, it would be an error to consider it as the only source of results for a real healthcare analysis. To address this shortfall, a more complex strategy should be designed, combining a set of techniques and methods (e.g., process mining, data mining [21,30]), and also combining contextual elements such as legal context, clinical practice guidelines, computer-aided policies and knowledge representation [31–33].

Other approaches have analyzed the organizational perspective and the collaboration in the medical centers through HIS in different ways. Frameworks have been defined to identify collaboration in clinical contexts through significant aspects such as coordination, communication and the shared information. Content analysis has studied factors such as redundancy, related information, ambiguity and compatibility, to understand how these mechanisms influence collaboration [34]. Principles on how to integrate professional quality systems, delegation of tasks, integrated planning, and process-aware information technologies have been studied, so as to improve the quality of the collaboration in healthcare programs [35]. Studies have gathered information on the structure of human activity systems and how these can be supported through medical records to implement innovative collaborative solutions [36]. Definitions of how process-aware information systems can be designed and implemented have been described in the past, being this a necessary and significant aspect in the development and analysis of collaboration in a HIS [16]. And finally, evaluation of the design and implementation of HIS has been carried out to identify challenges on how HIS can support, in the most efficient way, the professionals working in the healthcare domain [37].

2.5. Additional data science fields

Additional to the process mining field, other fields have had significant usage and relevance when used with emergency room data, providing not only knowledge about the executed process but providing tools to analyze data, predict outputs, recommend tasks, or even provide guidance on how to act under certain circumstances.

Data Mining in the ER has provided insight on how to triage and classify more accurately patients in the ER under different conditions [38], determine the best treatment for patients [39], improve outcomes and reduce expenditures through better disease management [40], among others.

Machine Learning techniques have been successfully applied on predicting cardiac arrest in critically ill patients [41], predicting in-hospital mortality [42], reducing waiting times [43], improving quality of care [44], among others.

In general, data science techniques can help to describe the process, understand how it is executed, when it is executed, and to describe its main characteristics, in order to characterize and predict the outcomes in different scenarios.

3. Method

The Design Science Research paradigm [45] was followed to conduct this research. This paradigm includes the following phases: problem identification & motivation, objectives of a solution, design & development, demonstration, evaluation, and communication. Following the problem and motivation previously described, we continue with the design & development phase, where the method for discovering organizational patterns was created, including its different steps: extracting the data, building an event log, applying process mining techniques to obtain the role interaction models, and evaluating the models with clinical experts in the domain. For the demonstration phase, the method was applied to a real case study, so as to validate the applicability of the method for discovering organizational patterns in a real healthcare environment. This case study will be described in Section 4 and its results in Section 5. The evaluation was conducted with physicians and clinical experts of the hospital where the case study was carried out. The last phase of the Design Science Research paradigm, communication, takes place in academic and healthcare domain proceedings.

There are generic methodologies for applying data mining, such as the CRoss-Industry Standard Process for Data Mining (CRISP-DM) methodology [46]. CRISP-DM is a popular methodology used in the context of DM projects, which presents a non-rigid sequence of six phases that can be used in a real environment, facilitating the creation of models that can be used to support business decisions. There are also some generic methodologies for applying process mining [11]. However, the method proposed in this article is based on the question-driven methodology for analyzing ER processes using process mining proposed in [13]. Accordingly, the proposed method was adapted with the aim of discovering role interaction models among healthcare professionals that collaborate in ER processes by analyzing the data stored in hospital information systems using process mining techniques.

The proposed method consists of four steps: (1) to *extract* the required data, (2) to *build* the event log, (3) to apply *process mining* tools and techniques to establish the role interaction models, and (4) to *validate* the obtained results with ER experts. The steps of this method are shown in Fig. 1, and are explained in more detail below.

(1) **Data extraction.** In order to analyze a healthcare process, data from hospital information systems (HIS) needs to be extracted. The extracted database must contain clinical information about episodes, i.e., data related to patient diagnosis, treatment, prescribed medicines, and severity of the episode. The process mining discipline requires an event log that contains at least a case ID, the

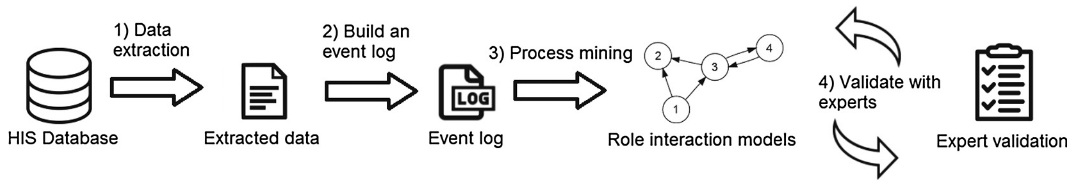


Fig. 1. Proposed method for discovering role interaction models.

performed activities, and a timestamp [11], although additional information may also be used, e.g. the resource executing or initiating the activity [47]. To discover role interaction models, our proposed method requires the extracted data to contain at least the following attributes: an episode identifier, the performed activities, a resource or performer and its role, and a timestamp to understand the order of occurrence of the events. In addition, the data must contain the diagnosis and the severity of the disease. To obtain this information, some queries to the HIS database must be executed. It is recommended to store the resulting tables as a CSV (comma-separated values) file for easy access and manipulation.

The main challenge to extract data from HIS about how healthcare processes have been performed, is that the information stored in the HIS is not process-analysis friendly. Data are split into many tables and the sequence of the activities cannot be determined easily, so it is necessary to make several queries to obtain the aforementioned information. Therefore, several scripts might be required to manipulate these tables and create an event log. Data can be extracted into structured data reference models that are process oriented and can help build the event log and provide higher quality data. A proposed data reference model was previously presented in [13] and can be used as an intermediate data model.

The granularity of the activities is usually very detailed, so it is also necessary to group low level activities to create high level ones. Since the objective is identify role interaction models, only activities performed by the same resource will be grouped. High level activities must have a name that describes all the activities contained within it, e.g., measuring body temperature, blood pressure, pulse rate and oxygen blood level might all be grouped into the "*Vital signs*" activity. The number of high level activities will depend on the properties of the process being analyzed.

(2) **Event Log Creation.** Data extracted from the HIS is known as raw data, because it is obtained directly from several distributed sources. The following task is to group the distributed raw data and create an event log that maintains consistency with the original data sources.

For example, in our case study, three scripts were created to manipulate the data and build an event log:

- The first script was used to associate activities, resources and their role, with the corresponding episode ID, in order to create events. The goal of this script is to transform raw data into table oriented data, which are organized in a structured format.
- An intermediate was used to extend the previously generated events by adding information about physicians, nurses and other healthcare professionals that participate in each episode. For each event, the script adds a label that distinguishes which of all physicians or nurses that appear in that episode performs each activity.
- Finally, the last script was used to read the information of all events and transform it into an event log, which is finally written to a CSV file.

  The process to obtain the resource level event log is shown in Fig. 2a. This event log contains information about each performed activity, its performer, the performer role, a timestamp, and other clinically relevant data, such as, episode severity, diagnosis, and discharge destination.

  An event of this log includes, for example, the information presented in Table 1. This kind of event log is recommended for analyzing the interaction between roles at a high level.

  The CSV file becomes the event log at a role-resource level if the intermediate script is used. Fig. 2b shows the process to obtain the event log at a role-resource level. This event log contains the information in the event log at a resource level, and it also contains an extra field that distinguishes the different resources of the same role that participate in the same episode. An event of this log is shown in Table 2. In this example, the first nurse appearing in the episode executed the "*Vital Signs*" activity. This kind of event log is recommended for analyzing the interaction between distinct resources belonging to the same role in a given episode, and also the interaction between resources from different roles.

(3) **Process mining.** This step includes two main tasks. The first one corresponds to the usage of process mining filtering tools in order to check data quality, remove data noise and eliminate incomplete episodes. The second task is the application of process mining algorithms on the filtered event log in order to discover and analyze role and role-resource interaction models.

In order to identify quality issues and address them, several guidelines have been considered when executing these filters [17,48,49]. For example, in our case study, in order to perform the first task it was necessary to apply two distinct filters:

- Completeness Filter: This filter was applied in order to eliminate noise and only keep episodes with complete information. The criterion used was the completeness of the episodes, considering only episodes with information about severity, diagnosis, discharge destination, and professionals who treated the patient. Episodes that did not have one or more of these fields were removed from the analysis. About 6% of the episodes were removed after applying this filter.
- Clinical Relevance Filter: Once the episodes with the required information were identified, the second filter was applied in order to obtain episodes that finished with a clinically relevant discharge, and to eliminate episodes in which the patient refuses to receive care, or is discharged home without any further treatment. About 12% of the episodes were removed after applying this filter.

  For this step, Disco[3] tool for process mining is used. This tool includes filtering capabilities to explore the data and remove cases and events that do not match with the filtering criteria defined above. Disco also includes a discovery algorithm inspired on the Fuzzy Miner algorithm [50], which can be use to generate the interaction models. Disco generates a diagram like the one presented in Fig. 3, where nodes represent the resources and the arrows represent a sequence of activities carried out by two resources. An arrow from node A to node B indicates that an activity performed by A is immediately followed by an activity performed by B. This sequencing of activities also describes the direction of the interaction between the two resources. The node color and the arrow thickness indicate the absolute frequency in all the event log. Darker nodes are roles that appear more frequently in the episodes, while lighter nodes appear with less

[3] https://fluxicon.com/disco.

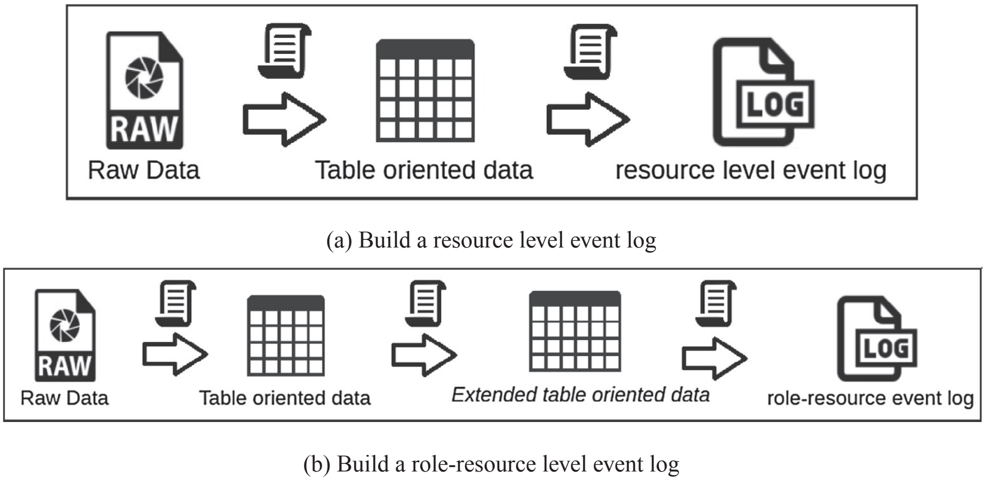


Fig. 2. Steps for building an event log.

Table 1
Information included in an Event log.

Episode Id	1
Activity	Examination
Resource	R1
Role	Physician
Specialist	Pediatrician
Timestamp	01-07-14 10:10
Diagnosis	Appendicitis
Triage color	Green
Discharge	Hospitalization

Table 2
Information included in an Event log at resource Level.

Episode Id	1
Activity	Vital Signs
Resource	R1
Role	Nurse
Role-resource	Nurse 1
Timestamp	01-07-14 10:15
Triage color	Green
Diagnosis	Appendicitis
Discharge	Hospitalization

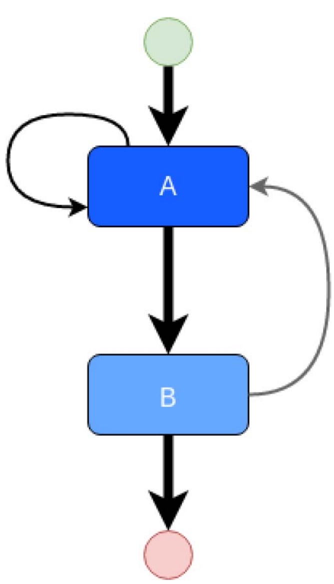


Fig. 3. Representation of a role interaction model in Disco.

frequency. In the same way, a thick arrow shows that a sequence of activities performed by two resources, one after the other, is frequent, while a thin arrow shows a very uncommon sequencing of activities performed by two resources. An incoming arrow from the initial circle to a given node indicates that the resource starts some of the episodes, and an outgoing arrow from a given node to the end circle indicates that the resource performs the last activity in some episodes.

After filtering, a discovery algorithm to analyze the event log was used. The result is a role interaction model. These models could have different characteristics depending on the severity of the episodes or the diagnosis.

(4) **Expert Validation**. After the role interaction models have been obtained in step 3, it is important to validate them with ER experts in order to obtain their feedback. This validation will be useful to verify if the results correspond to similar behavior or significant insights about the reality of ER processes. It is recommended to perform this validation with at least one expert or process owner that has the knowledge of the ER processes, the ER infrastructure, and the actual needs.

First of all, it is advisable to prepare the instruments to carry out this validation process. It is recommend establishing a series of open and closed questions in a questionnaire or interview to obtain the experts' insight [51]. The five main questions that must be included are:

– What is the actual organizational process and how similar is it to the obtained results?
– What information does each model provides, and to what extent is it important for the performance of the ER processes?
– Are the identified roles the expected ones?
– Do role interaction models reflect the reality of the ER processes?
– Can these results help to improve the teams working in the ER processes?

The interview should be performed using these questions. A presentation should be done to the experts in which the resulting role interaction models, the assumptions made for the analysis (e.g. included or excluded activities), and the final findings are displayed to start the discussion.

The proposed method helped us to analyze role interaction models in our case study. The details of the case study are presented in Section 4; the results will be described in Section 5; meanwhile, Section 6 focuses on the discussion of the results.

4. Case study

The research on role interaction models led to a case study on a real

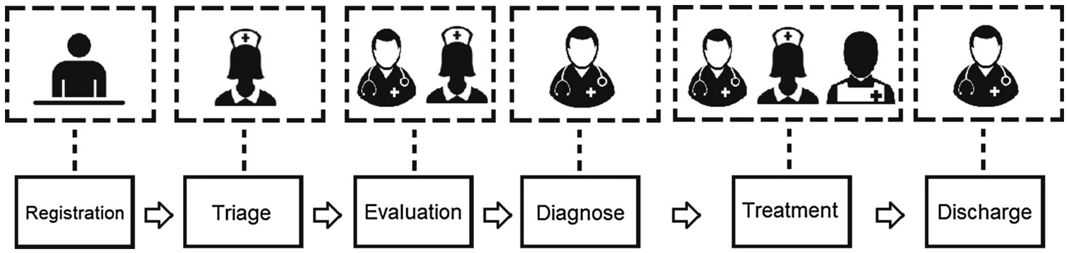


Fig. 4. ER process model.

scenario, which is aimed at seeing whether it is possible to detect any patterns among the resources that perform activities in an ER process by applying the method proposed in Section 3. The details of the case study are described in this section and the results are shown in Section 5.

4.1. Hypothesis

It is proposed to validate two hypotheses in the case study: (i) it is possible to find role interaction models and role-resource interaction models in ER episodes through process mining, and (ii) it is possible to describe role interaction models and role-resource interaction models, and to distinguish them by using specific criteria, such as, diagnosis or severity of patients.

4.2. Context

The case study is based on an Emergency Room process of a 500-bed, high-complexity university hospital in Santiago, Chile. It provides takes care of approximately 70,000 patients per year, 13% of them are hospitalized. The Emergency Room had been using their Electronic Health Record (EHR) for 18 months and it includes full clinical documentation for physicians and nurses, Computerized Physician Order Entry (CPOE) and integration with laboratory, radiology and billing systems. The goal of this ER is to screen, examine and provide care to patients who come without prior appointment, presenting a broad spectrum of illnesses and injuries, of which only a few may be life-threatening and require immediate care. A permanent challenge is to improve the quality of the service level provided, reduce overcrowding, and provide prompt and efficient care.

4.3. Process description

The research described in this paper is based on a case study related to the ER described in Section 4.2. Fig. 4 shows a reference diagram of the ER process followed by the hospital. The process begins when a patient comes to the ER. The receptionist registers the patient and the patient goes to a waiting room. Then, a nurse interviews the patient so as to perform the triage and determine the severity of the episode. The nurse is not only dedicated to this task, and can vary from episode to episode. After the severity is established, the patient returns to the waiting room until he/she is called to a cubicle for attention, where a professional measures his/her vital signs. The results are given to a physician, who becomes responsible for the episode. The physician evaluates the patient, asks the patient to undergo some examinations, diagnoses the patient, and then discharges the patient to his home or decides the patient needs to be hospitalized. Sometimes the patient leaves before treatment has started or explicitly asks to be discharged. In the ER studied, there are some medical guidelines that determine how the process should be executed. These guidelines should also have a major influence on how the interaction among the healthcare professionals take place, and thus on how the interaction is registered in the HIS. However, in a previous work [13], it was established that the actual process execution do not strictly follow these pre-defined guidelines.

The application of the steps of the proposed method in the case study are presented next.

4.4. Materials and methods

To carry on the case study the method presented in Section 3 was followed. The steps Data extraction and Event Log creation are specified as follows. The Results and Validation steps are presented in Sections 5 and 6, respectively.

4.4.1. Data extraction

Data were obtained from *Alert Data Warehouse Phase I*,[4] which corresponds to the HIS used in the ER of the hospital. The data are distributed between several tables and contained information about clinical episodes that occurred between July 1, 2014 and July 31, 2014. The data correspond to 7160 episodes, including 309,796 performed activities. The information is structured in a series of tables in CSV format. In those tables detailed information is found about the patient diagnosis, vital sign measurements, allergies, referrals, patient transportation, responsibility transfers, diagnosis, professional activities, drugs, discharges and triage data, and the date and time of execution.

Finally, it is decided only to use the following tables to build the event log for the organizational analysis:

- "*Professional activities report*": this table contains information about all the activities performed by all the resources involved in the episodes, including the episode ID, name of the activity, date and time, resource identification and specialty of the performer. This table provides most of the data needed to create the resource level event log and represents 56.7% of the total data used.
- "*Episode report*": this table adds information regarding the severity of the episodes. This represents about 6.1% of the total data used.
- "*Diagnoses report*": this table adds information about the diagnoses of diseases for each episode. This represents about 3.1% of the total data used.
- "*Activity detail report*": this table contains detailed information about clinically relevant activities: triage, vital signs measurements, taking note of patient complaints, performing diagnosis, referrals and discharges. It also contains the date and time of the activity, performer, and specialty of the performer. This table provides the data needed to create the role-resource level event log and represents 34% of the total data used.

4.4.2. Event log creation

For the role interaction analysis, an event log was built using the "*Professional activities report*", which contains the activities carried out by the staff of the emergency room, indicating also the date and time of completion of the activities, the resource identification, the role of each resource, physician specialty (if applicable), the diagnosis, and an episode ID. With the created scripts this table was read, the information

[4] http://www.alert-online.com.

Table 3
Example of an event log at resource level.

Episode Id	Activity	Resource	Role	Specialist	Timestamp	Diagnosis	Triage color
1	Triage	R1	Nurse	–	04/07/14 10:00	Appendicitis	Green
1	Vital Signs	R1	Nurse	–	04/07/14 10:10	Appendicitis	Green
1	Examine	R1	Nurse	Pediatrician	04/07/14 10:30	Appendicitis	Green

extracted, and grouped them by episode ID in a new file called "*episode information*". After that, the "*episode information*" file was read and a CSV file was created, called "*general roles*" that became the resource level event log. The columns that the event log contains are the following:

- *Episode ID:* the episode identifier.
- *Activity:* the task performed by a resource.
- *Resource:* the person that executes a specific task.
- *Role:* the role of the resource.
- *Specialist:* physician specialty (if applicable).
- *Timestamp:* date and time of the performed task
- *Diagnosis:* the diagnosis made by the physician, obtained from the "*Diagnosis report*".
- *Triage Color:* the severity of the episode, obtained from the "*Episode report*".

Table 3 shows an extract of the "*general roles*" event log. As can be seen, there are different types of activities, such as triage, vital signs measurements, examinations, taking note of patient complaints, performing diagnoses, referrals, discharges, among others. Originally, this event log contained 6760 episodes from the 7160 episodes recorded in the HIS. The reduction in the number of episodes was due to the fact that not all the episodes contained the minimum data to transform them into a trace of the event log. Some episodes did not have information about the activities carried out and/or about the specialists who performed those activities. The event log, after applying the completeness and clinical relevance filters, contains the following information:

- 5542 episodes, with 2029 variants among the different roles, and 5362 among the different resources.
- 21 different activities.
- 240 different resources, of whom 47 start the episodes, and 197 end the episodes.
- 697 different diagnoses.
- The mean duration of the episodes is 5.1 h.

In the "*general roles*" event log, four roles were discovered: physician, nurse, technician and medical assistant. This log will provide the general role interaction model, but it is important to perform a more detailed analysis that involves the interaction of resources of the same role in order to discover role-resource interaction models at a lower level. To do that, a second event log called "*physician and nurse*" with the goal of doing a deeper analysis was created, focusing on the distinct physicians and nurses interacting in the episodes. In the "*physician and nurse*" event log, focused only on physicians and nurses, professionals who belong to the same role and are involved in the same episode are also identified. To build the second log, the "*Activity detail report*" was used, considering only the clinically relevant activities aforementioned. The created scripts were used for the extraction of information. The intermediate script was used to count and distinguish physicians and nurses in each episode. This makes it easier to analyze the collaboration between people who perform the same role. Finally, the event log named "*physician and nurse*" was obtained, that becomes the event log at a role-resource level. An extract of the "*physician and nurse*" event log is shown in Table 4. This event log contained originally 6219 episodes from the 7160 episodes recorded in the HIS, since some episodes did not have information about the activities carried out and/or about the different specialists who performed those activities. After applying the completeness and clinical relevance filters, the event log contains the following information:

- 5175 episodes, with 69 variants.
- 10 different activities.
- 149 different resources, of whom 36 start the episodes, and 96 end the episodes.
- 6 different role-resources combinations.
- 690 different diagnoses.
- The mean duration of the episodes is 3 h.

The "*general roles*" and "*physician and nurse*" event logs were imported to Disco. The two filters described in the process mining step of the proposed method were applied to both event logs, with the objective of working only with episodes with complete and clinically relevant information for the doctors in charge of the ER process. Disco's discovery algorithm was applied to both event logs and found role interaction models and role-resource interaction models.

4.4.3. Mining tool

In the PM literature, there are several discovery algorithms [11] that can be considered as candidates to discover role interaction models. For the case study of this article, the Disco miner [52] was selected. Disco miner is a discovery algorithm implemented in the commercial tool named Disco (https://fluxicon.com/disco), which is based on the Fuzzy Miner algorithm [50]. Both of these algorithms take into account the frequencies of events and sequences as well, in order to create a process model (also known as a process map). The main reasons considered to choose this algorithm were its facilitates to handle complex processes, such as staff interactions in ER, and the possibility to obtain a process-map model as an outcome, model that can be easily understood and manipulated by healthcare professionals with little experience in the PM discipline. Furthermore, Disco facilitates the process analysis from different perspectives due to its event log filtering functionalities, which allow the analysis of diverse scenarios considering distinct data attributes.

Following, Section 5 presents the results obtained. The last phase, *Validate with expert*, will be presented in Section 6, with the discussion of the results.

5. Results

For the case study two analyses were performed: first, an analysis of the interaction among roles using the "*general roles*" event log to obtain role interaction models; and second, an analysis of the interaction among different resources performing the physician and nurse roles during each episodes using the "*physician and nurse*" event log to obtain role-resource interaction models.

For both aforementioned analyses, first the general model was explored, considering all events, and then a study was performed at two different sub levels in order to describe more specific interaction models and compare them with the general interaction model.

The first sub level is at the triage color level. The triage information of the episodes was used, according to the Manchester triage system, which is a priority classification system based on expert knowledge. It

Table 4
Example of an event log at role-resource level.

Episode Id	Activity	Resource	Role	Role Tag	Specialist	Timestamp	Diagnosis	Triage color
1	Triage	R1	Nurse	Nurse 1	–	04/07/14 10:00	Appendicitis	Yellow
1	Vital Signs	R1	Nurse	Nurse 1	–	04/07/14 10:10	Appendicitis	Yellow
1	Examine	R2	Physician	Physician 1	Pediatrician	04/07/14 10:30	Appendicitis	Yellow
1	Biometrics	R3	Nurse	Nurse 2	–	04/07/14 11:00	Appendicitis	Yellow

categorizes patients into five different colors (red, orange, yellow, green and blue), depending on the severity of the episode and the immediate need of care. Red represents the most severe episodes; green, the lowest severity cases; and blue, the non-urgent episodes. Red episodes will have higher priority and shorter waiting times, while blue episodes will have to wait longer [53].

The second sub level is at the diagnosis level. A diagnosis is the determination of the cause of a patient's illness or suffering by the combined use of physical examination, patient interview, laboratory tests, review of the patient's medical records, knowledge of the cause of observed signs and symptoms, and differential elimination of similar possible causes [54].

These two characteristics were selected because they are the most important in an episode, according to ER experts.

The discovery of role interaction models was performed by importing the "*general roles*" event log obtained in the "*build an event log*" step of the method, to the Disco software. As a first step for importing an event log to Disco, it is necessary to specify which column will be chosen as the case ID, the activity or task performed, the resource and the timestamps. For example, the episode ID is specified as Disco's case ID. Notice that it is necessary to use Disco in a non-conventional manner in order to discover role interaction models: specifying the roles of the resources as Disco's activities, so that nodes in Disco's diagrams represent roles instead of tasks. Fig. 5 shows the general role interaction model, considering all the 5542 episodes. The nodes represent the roles, and the arrows represent the sequencing of activities performed by two roles (or the direction of the interaction between two resources).

To remove less relevant interactions, filters were applied to the episodes so as to obtain a simplified role interaction model, and to focus the analysis on patterns that repeat across several episodes. First, eliminating the least frequent interactions between two roles: when its frequency was less than 10% of the total episodes. Then, repeating the same filtering process until all interactions with low frequency were

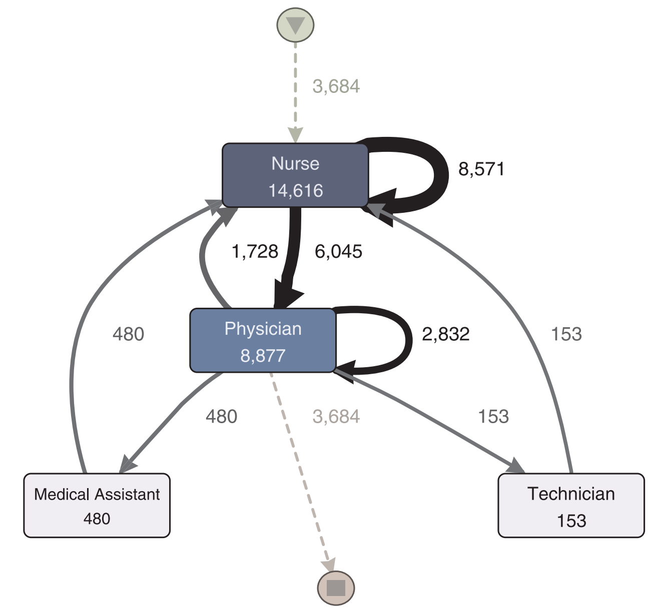


Fig. 6. General role interaction model after filtering less frequent interactions: the "*lung model*".

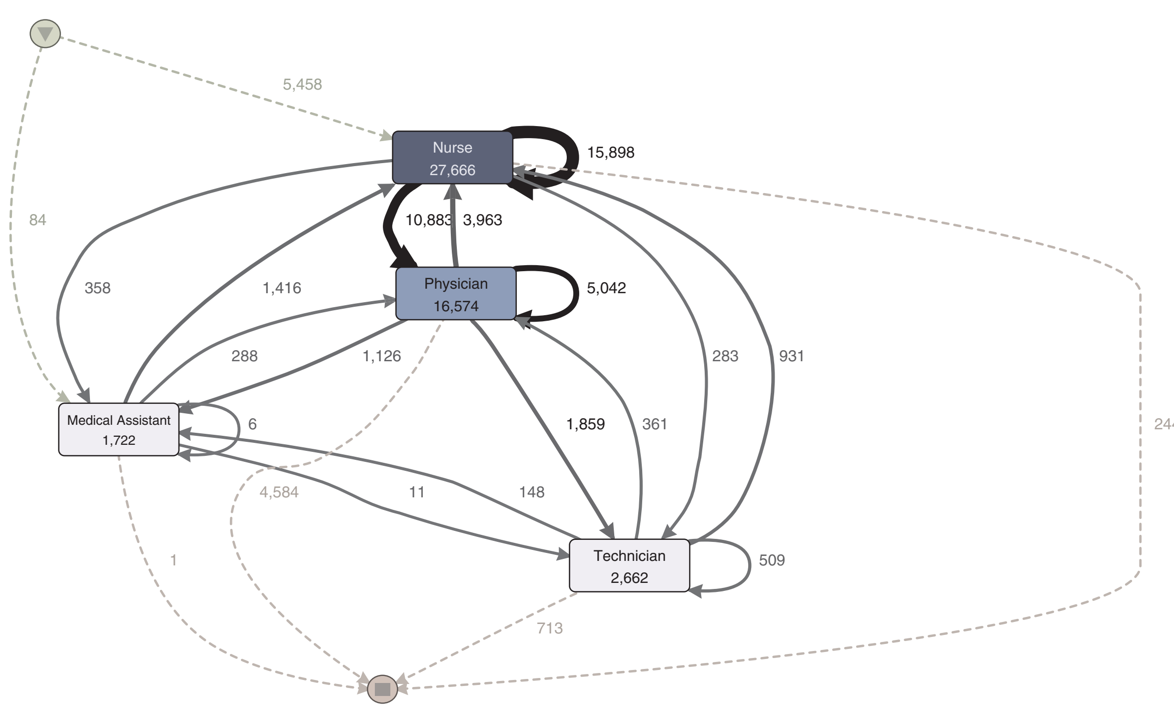


Fig. 5. General role interaction model, considering all episodes.

Table 5
Frequency of participation of each healthcare role in the episodes of the "*lung model*".

	Case Frequency
Nurse	3684 (100.0%)
Physician	3684 (100.0%)
Medical Assistant	462 (12.5%)
Technician	136 (3.7%)

eliminated. Given the complexity of ER processes, and that the event log presents many variants, it is expected that after applying the mentioned filter, several of the episodes corresponding to the least frequent variants of the process have been removed.

Fig. 6 shows the general role interaction model after filtering those episodes that contain less frequent interactions, leaving only 54% of the episodes. This model is called the "*lung model*" because the shape of the obtained model looks like a lung. The "*lung model*" describes how the roles interact with each other in most of the episodes. In this model, four roles were found: physicians, nurses, technicians and medical assistants. Physicians and nurses appear in 100% of the episodes, while medical assistants appear in 12.5% of the episodes and technicians appear only in 3.7% of the episodes (see Table 5). The nurse is the role that performs more activities (in average, 3.97 activities per episode), followed by the physician (in average, 2.41 activities per episode), as shown in Table 7. On the other hand, the technician and the medical assistant participate less often in the episodes (in average, 0.04 and 0.13 activities per episode, respectively). There are seven interactions among the roles. The most important ones are the thickest arrows: from nurse to nurse, from nurse to physician, from physician to physician, and from physician to nurse. The other interactions are less frequent: physician to medical assistant, medical assistant to nurse, physician to technician, and technician to nurse.

After the discovery of the "*lung model*, an analysis was performed at the triage color sub level. Fig. 7 shows the role interaction models obtained for different triage colors. Blue and red episodes represent less than 2% of the episodes (see Table 6). Green episodes are 50.1%, yellow episodes are 39.0%, and orange episodes are 9.7% of the episodes. Green, yellow, orange and red episodes have similar role interaction models; they also display the "*lung model*". Meanwhile, the role interaction model for the blue episodes is different from the others since technicians do not participate in these episodes.

Table 7 shows the number of times each healthcare role participates in an episode by triage color. The participation of almost all healthcare roles increases as patients' priority increases, with the exception of the medical assistants. Fig. 8 shows how the participation of both nurses and physicians increases as patients' priority increases and how nurses' participation growth faster than physicians' participation. Fig. 9 shows how the medical assistants participate more often in the yellow and orange episodes. Meanwhile, technicians have a greater participation in the orange and red episodes, and are practically not required in blue and green episodes.

The study at the diagnosis sub level was also performed. Fig. 10 shows the role interaction models for the four most common diagnosis (see Table 6) in the event log: Flu, which represents 9.7% of the episodes; Cold, representing 6.1% of the episodes; Abdominal Pain, which represents 4.3% of the episodes; and Headache, representing 3.6% of the episodes. The role interaction models for these diagnosis are similar, but technicians do not participate in episodes of patients diagnosed with Cold. Variations of the general pattern were found in other diagnosis that are less common. ER professionals work in a standardized way for common diseases, while in other episodes they appear to work following other patterns. For example, Fig. 11a shows the role interaction model for Contusion episodes. This diagnosis is present in 1.8% of the episodes. In the Contusion role interaction model, there is no technician involve in the treatment of patients, and physicians interact more often with themselves. Fig. 11b shows another example, a role interaction model for Burn episodes, which represent 0.5% of the episodes. In Burn episodes, only physicians and nurses participate.

Table 8 shows the number of times each healthcare role participates in an episode by diagnosis. Three groups can be identified. Group 1 includes those diagnosis in which most healthcare roles participate less often than in the general case. Notice that in this group, technicians do not participate in Cold episodes, and neither technicians nor medical assistant participate in Burn episodes. On the other hand, Group 2 includes those diagnosis in which most healthcare roles participate more often than in the general case. Group 3 includes two diagnosis that have a singular behaviour: Appendicitis and Contusions. In Appendicitis episodes, all roles participate significantly more than in the general case: about twice in the case of nurses and physicians, and even more, in the case of technicians and medical assistants. Meanwhile, in Contusion episodes, nurses participate less often, technicians are not involved, and physicians are assisted by medical assistants in order to treat the patients.

After the first analysis of the "*general roles*" event log, the second analysis was performed at the role-resource level by importing the "*physician and nurse*" event log to Disco. Fig. 12 shows the role-resource interaction model obtained by using this event log, which distinguishes the different nurses and physicians participating in the episodes by a label with a number next to their role (e.g., Nurse 1, Physician 1, Nurse 2). In this case, 5% of less frequent interactions (arcs) were filtered. For most episodes, there is one physician and two different nurses participating, and in some episodes, there is a third nurse or a second physician involved. As in the "*lung model*", a nurse (Nurse 1) opens the episode and a physician (Physician 1) closes it. The main sequencing of activities here are: Nurse 1 to Physician 1, Physician 1 to Nurse 2, and Nurse 2 to Physician 1. All episodes begin with a nurse (Nurse 1), who performs the triage to the patient. Then, a physician (Physician 1) or another nurse (Nurse 2) performs another activity, after Nurse 1. Notice that is uncommon for Physician 1 to interact with the same nurse who performed the triage. The sequencing of interactions can be interpreted as follows. Physician 1 acts as the person responsible for the episode, transferring the work to Nurse 2 or to him/herself. Nurse 2 transfers the work back to Physician 1, or requests the collaboration of another nurse (Nurse 3). In the latter case, Nurse 3 always transfers the work back to Physician 1. Physician 1 may also refer the patient to another physician (Physician 2), who transfers the work to any of the duo who is heading the episode (Physician 1 or Nurse 1). All episodes are ended by Physician 1. Notice that in the event log there is no evidence of transferring of work, but rather of the sequencing of activities performed by the different resources.

As previously done with the "*general roles*" event log, the role-resource interaction model filtered by triage color, and then by diagnosis were studied. Fig. 13 shows the role-resource interaction model for different triage colors. Although the layout looks different, the same role-resource interaction patterns observed in the general model occur for most triage colors. Red episodes are the exception, where only one physician and one nurse take care of the patient.

In the role interaction models obtained with the "*general roles*" event log filtering by diagnosis, no significant differences with the "*lung model*" were discovered. In contrast, when analyzing the "*Physician and nurse*" event log filtering by diagnosis, some variations were observed to the general role-resource interaction model. Three different role-resource interaction patterns were found.

Fig. 14 shows the role-resource interaction model for two very common diagnosis, Flu and Headache, in which Physician 1 requests the collaboration of a second physician (Physician 2) more frequently than in the general role-resource interaction model. Although the interaction occurs through a referral, Physician 1 seems to be the responsible of the episode, and therefore the work is always transferred back to him. Accordingly, Physician 2 transfers the work back to

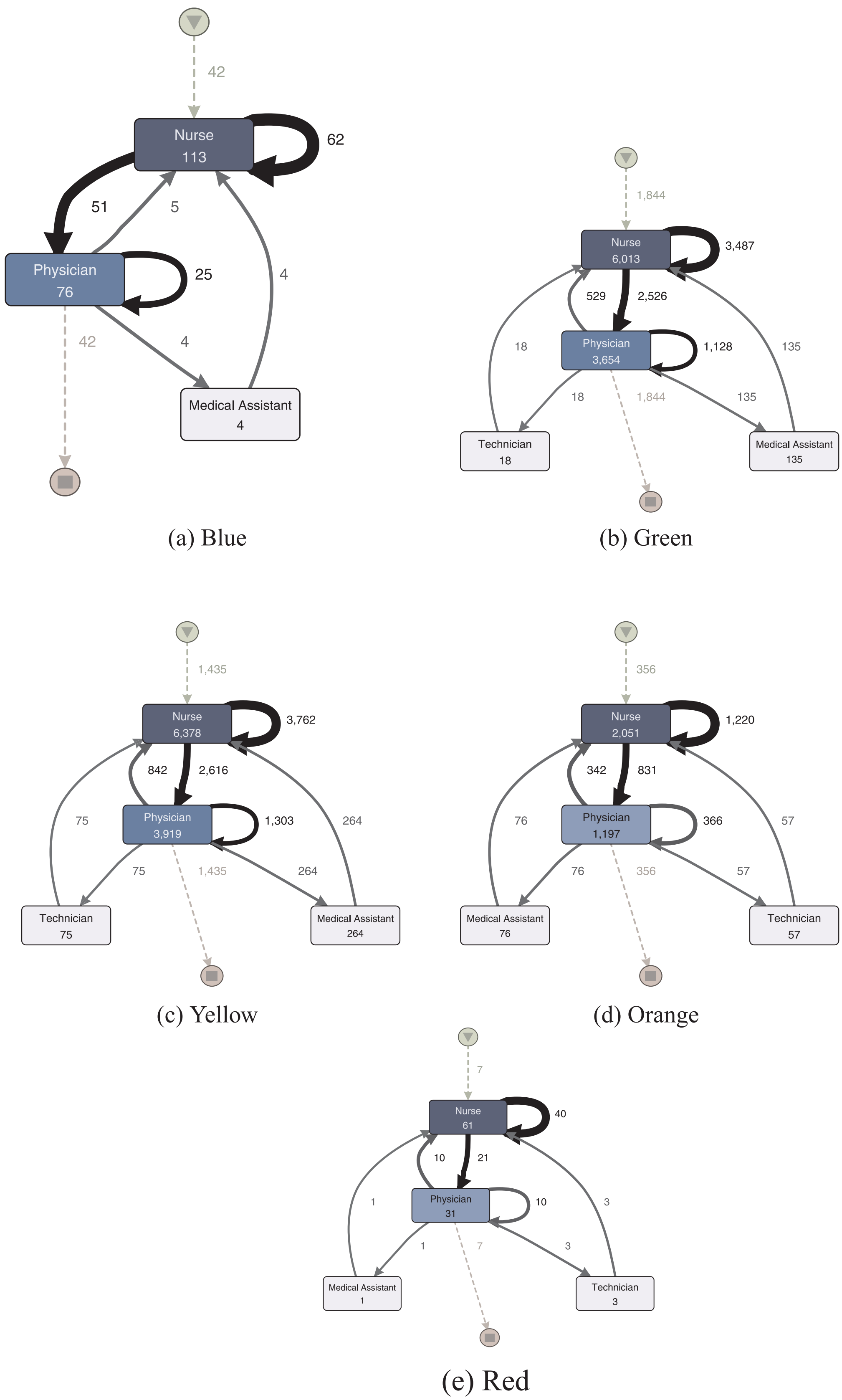


Fig. 7. Role interaction models filtered by triage color.

Physician 1 or incorporates Nurse 2. Notice that Physician 1 never asks the collaboration of Nurse 2 directly. Notice also that there seldom is a loop from Physician 1 to him/herself.

Fig. 15 shows the role-resource interaction model for other two very common diagnosis, Abdominal pain and Cold. A strong collaboration can be seen between Physician 1 and Nurse 2. Nurse 3 is only incorporated in a few episodes. Notice that the participation of a second physician is not required.

Table 6
Frequency of episodes of the "*lung model*" filtered by triage color and diagnosis.

	Case frequency
Blue	42 (1.1%)
Green	1844 (50.1%)
Yellow	1435 (39.0%)
Orange	356 (9.7%)
Red	7 (0.2%)
Flu	359 (9.7%)
Cold	226 (6.1%)
Gastroenteritis	187 (5.1%)
Abdominal pain	158 (4.3%)
Headache	132 (3.6%)
Contusions	68 (1.8%)
Burns	17 (0.5%)
Appendicitis	6 (0.2%)

Table 7
Ratio of participation of each healthcare role per episode by triage color.

Triage color	Nurse	Physician	Technician	Medical Assistant
General	**3.97**	**2.41**	**0.04**	**0.13**
Blue	2.69	1.81	0.00	0.10
Green	3.26	1.98	0.01	0.07
Yellow	4.44	2.73	0.05	0.18
Orange	5.76	3.36	0.16	0.21
Red	8.71	4.43	0.43	0.14

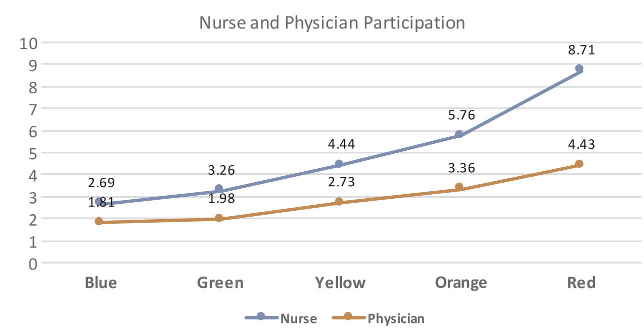


Fig. 8. Ratio of participation of Nurses and Physicians per episode by triage color.

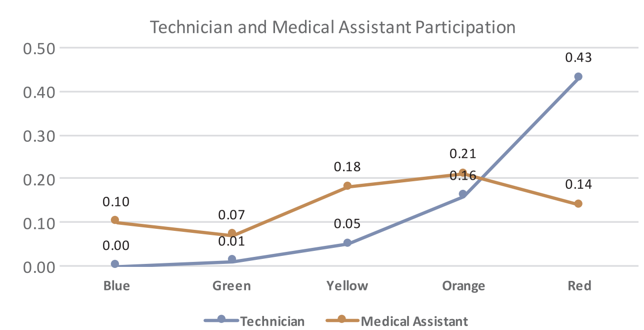


Fig. 9. Ratio of participation of Technicians and Medical Assistants per episode by triage color.

Finally, Fig. 16 shows the role-resource interaction model for two diagnosis, Contusions and Burns, in which the treatment is handled by only two professionals: Physician 1 and Nurse 2.

Process mining tools and techniques can help to obtain and understand different interaction patterns. These patterns can be analyzed at a high level considering all the participating roles, and at a detailed level, focusing on the most important roles and distinguishing the different professionals performing each role in the episodes. The interaction models can also be characterized by triage color or diagnosis, finding different interaction models for different episodes' severity and different episodes' diagnosis. Some common elements can be found among all interaction models: nurses and physicians are present in all episodes, nurses always open the episodes, and physicians always close them. The obtained interaction models were analyzed with expert physicians in order to validate them and contrast them with the actual behavior. The analysis of the results is presented in the next section.

6. Discussion

In this section, after answering the research question and confirming that role interaction models in ER can be discovered through process mining techniques, the obtained results, shown in Section 5, are discussed. The obtained results are discussed from the healthcare perspective, analyzing the relevance of the obtained role interaction models to understand how healthcare professionals collaborate in ER processes.

The case study was analyzed with an expert in ER, a physician who has an overview perspective of the ER needs and understands the clinical characteristics of the ER processes. During this task, the five basic questions specified in Section 3 were considered, as they correspond to the main validations needed. After the case study is discussed, the main limitations are presented.

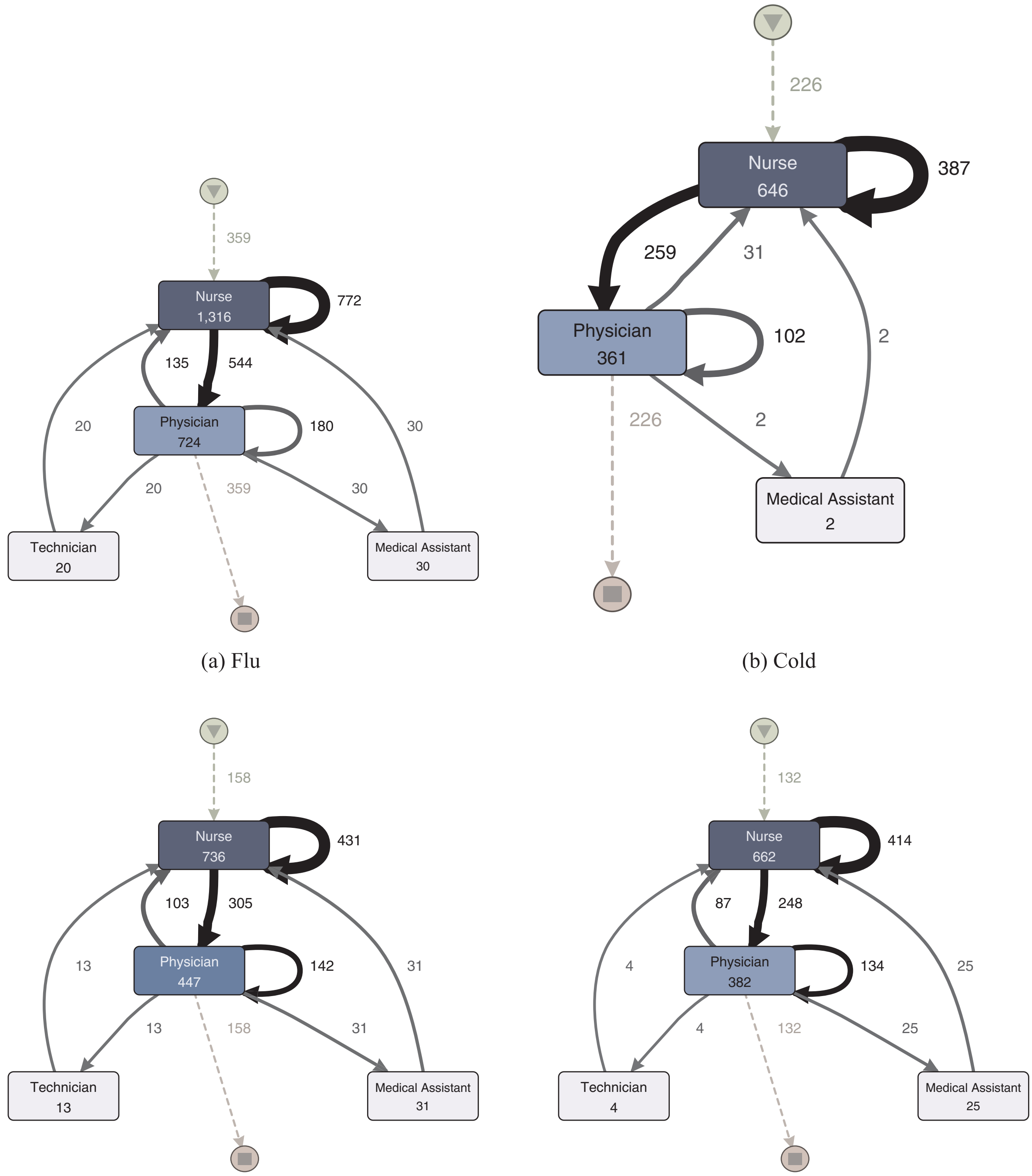
6.1. Case study

The general role interaction model reflects the overall execution of an episode in the ER. The ER and Internal Medicine Expert confirmed that the "*lung model*" is followed in most episodes and describes the most common role interaction patterns (see Fig. 6). In this model, nurses start most episodes and physicians close most of them. As expected by the ER expert, the interaction is more intense among nurses and physicians. The physician is in charge of the episode and interacts with other roles through ordering tests or treatments, as well as referrals to other physicians. The other two roles, technicians and medical assistants, are support roles, perform less activities, and therefore their interactions are less intense. Although one could expect differences in interaction models in patients with different acuity levels, these differences may emerge from other attributes not captured in the role interaction model such as the type and timeliness of those interactions.

The expert considered that the possibility of filtering the episodes according to different episode attributes is potentially useful. The fact technicians do not participate in blue episodes is expected since usually less severe episodes do not require further tests. Discovering different role interactions while dealing with different diagnoses is helpful because it can help trigger actions to improve and standardize the protocols for dealing with episodes corresponding to different diagnoses.

Another aspect mentioned by the experts is that the ER process of this case study were very flexible, because any nurse or physician could treat a patient and, therefore, it is a fact that almost any nurse might interact with any physician at least once. Therefore, any analysis at the resource level considering, for example, the name of the professional, would show a very complex model, not useful for a clear analysis. Consequently, the ER expert considered very interesting the role-resource interaction models because they allow a more insightful analysis at a reasonable abstract level.

In the role-resource interaction model shown in Fig. 12, it was found that a physician and two nurses work together in most episodes. The first nurse starts the episodes by doing the triage to the patient, and then delegates the job to the first physician, who together with the second nurse, treats the patient. Sometimes, a third nurse and a second physician might also be involved in the treatment. The first physician, as the healthcare professional who is responsible for the episode, usually closes the episode. This overall description of the interaction among healthcare professionals was confirmed by the ER expert. For



(a) Flu

(b) Cold

(c) Abdominal pain

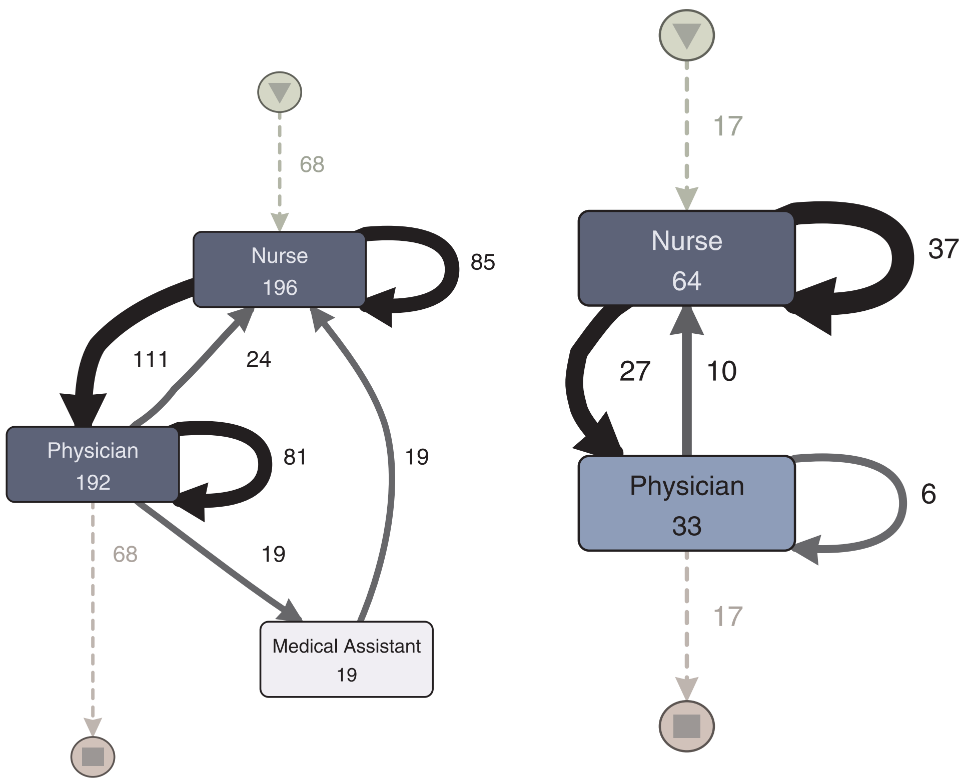
(d) Headache

Fig. 10. Role interaction models filtered by diagnosis.

example, it is a fact than when patients arrive, they are classified according to the triage system by a first nurse, and when they called for attention in the cubicle, the available physician and nurse will continue with the episode. Moreover, the role-resource interaction models allow to understand in more detail the interaction dynamics for episodes of different severity or diagnosis.

Analyzing the resulting models and information obtained with the ER and Internal Medicine Department experts, allowed to confirm that process mining techniques can help to understand how roles interact in ER episodes. With this, the first hypothesis has been validated: it is possible to find role interaction models and role-resource interaction models in ER episodes through process mining. The role interaction models for episodes filtered by triage color or diagnosis show interesting differences. As a consequence, the second hypothesis was also corroborated: it is possible to describe role interaction models and role-resource interaction models, and to distinguish them by using specific criteria, such as, diagnosis or severity of patients.

In general, the application of the proposed method in the case study provided several benefits. It generated new information about how the different roles interact, including the discovery of real-life patterns. It allows to understand what differences exist for episodes of different severity or diagnosis. Particularly, at the role-resource level, different interaction models were found for different severity and diagnosis. These models allow to discuss whether different interactions might be useful for providing a better treatment to the patients. The discussion, in turn, can trigger actions to improve and standardize the protocols to



(a) Contusions (b) Burns

Fig. 11. Role interaction models for contusions and burns episodes.

Table 8
Ratio of participation of each healthcare role per episode by Diagnosis.

Group	Diagnosis	Nurse	Physician	Technician	Medical Assistant
	General	**3.97**	**2.41**	**0.04**	**0.13**
Group 1	Cold	2.86	1.60	0.00	0.01
	Flu	3.67	2.02	0.06	0.08
	Burns	3.76	1.94	0.00	0.00
	Gastroenteritis	3.87	2.31	0.02	0.03
Group 2	Abdominal Pain	4.66	2.83	0.08	0.20
	Headache	5.02	2.89	0.03	0.19
Group 3	Appendicitis	7.67	4.83	0.17	0.83
	Contusions	2.88	2.82	0.00	0.28

deal with different types of episodes.

In order to apply the method in additional settings where interaction patterns are relevant, several requirements can be identified: data from a HIS must be available for analysis, the activities stored in the HIS must reflect the executed process (e.g., by registering as many activities performed in the ER processes as possible), and it must be structured in a process-oriented way, e.g., by using a data reference model to help to structure data, such as the one we proposed in [13].

6.2. Limitations

Process mining has been applied in more than 70 case studies in healthcare [11,17,21,55], providing significant insights about the executed processes, identifying the executed activities, the order in which they are performed, their execution time, the activities that are considered bottlenecks and increment the episode duration times, among others. However, all process mining methods proposed, including the one proposed in this paper have two well-known limitations. First, process mining is a data-based analysis approach, i.e., it uses the event data recorded on the HIS to discover, verify, or analyze a hospital process. Any event not recorded in the system is considered invisible to the analysis; informal corridor discussions or non-recorded verbal orders are some examples in a hospital context. Some authors posit that this limitation may be mitigated by the use of process-aware information systems [56] (i.e., when there is an explicit definition of the process in the HIS), however, given the amount of variability in ER processes, this may not be feasible. Second, as a data-based analysis technique, process mining relies on the quality of the data, i.e., misrecorded or incomplete data could lead to biased or directly wrong analysis of the process [48]. For example, several medical interactions performed in different moments of time could have the same timestamp (typically every hour o'clock). In the literature, several works mitigate this limitation by using partial ordered event data [57], or with strategies for data repair based on a pattern repository [49] or reference model alignment [58].

7. Conclusions and future work

In this case study, a method based on process mining techniques in order to discover role interaction models that describe and match the collaboration among ER professionals who treat patients in an ER episode has been applied. Each step of the proposed method has been executed: obtaining the necessary data on ER episodes from a HIS system, building an event log, discovering role interaction models in the ER episodes using process mining techniques, and finally, validating the results obtained with an ER expert.

The main contributions of this research are: (i) the proposed method, as a guide for using process mining tools and techniques to help discovering role interaction models, and analyzing them; (ii) the use of process mining tools and techniques to discover both role interaction models and role-resource interaction models, and then characterize them, (iii) the use of filters based on ER episode attributes to analyze more in detail the role interaction models, and (iv) the application of the proposed method in a real case study with the purpose of

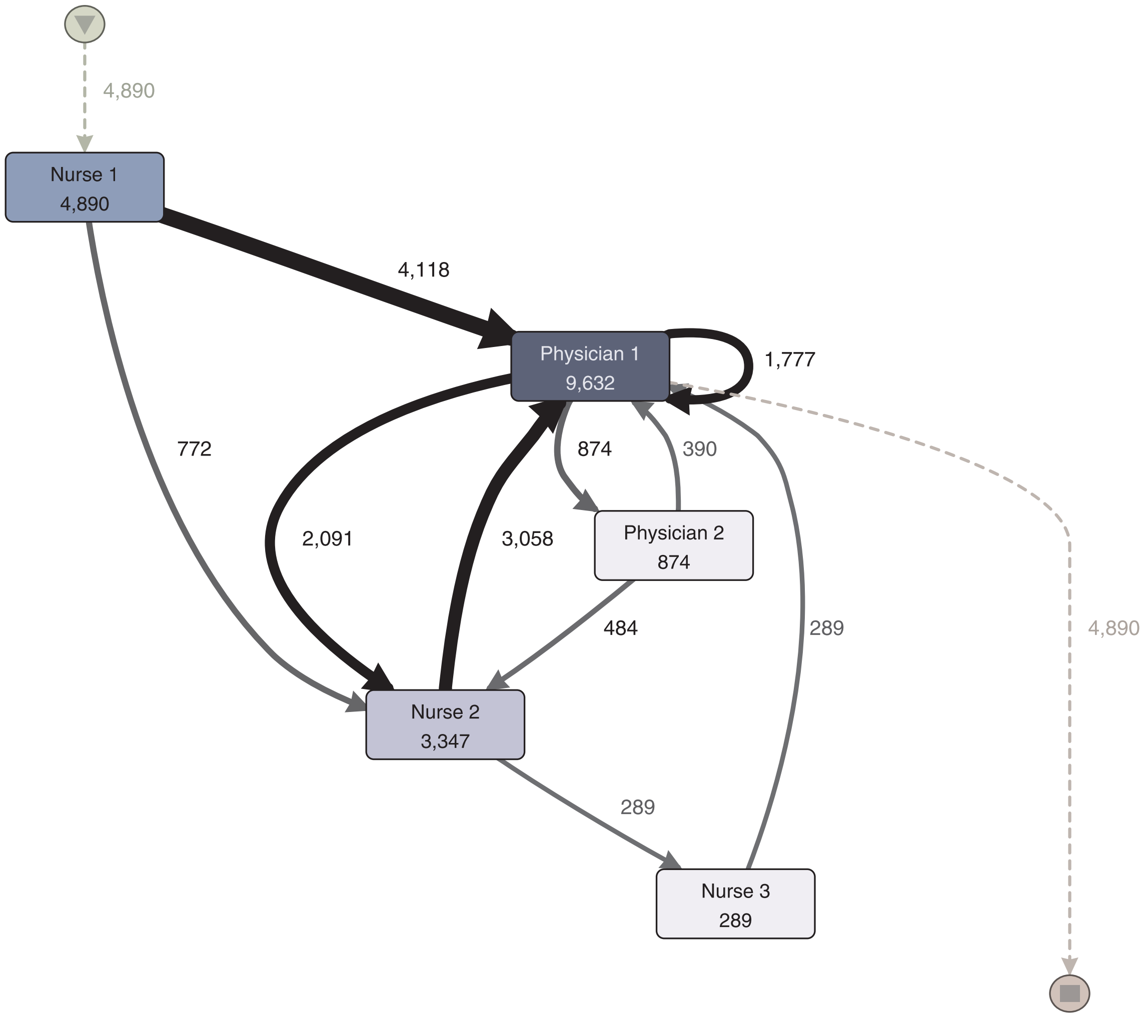


Fig. 12. General role-resource interaction model after filtering less frequent interactions.

illustrating its pertinence.

The discovery of role interaction models through process mining techniques shows another way of giving insight and visualizing collaboration patterns that describe professional interactions in ER processes, which match those observed by the own healthcare professionals. These results are achieved through the combination of process mining techniques with the healthcare domain knowledge, which provided significant insights for both process experts and healthcare professionals. This knowledge allows uncovering opportunities to improve ER processes from the organizational perspective.

The application of the proposed method in a real case study generated relevant insights about how healthcare professionals collaborate at the role level:

– A main role interaction pattern was identified (the lung model), where there is high interaction between a nurse and a physician, and less interaction with a medical assistant and a technician.
– Prevalence of this main role interaction pattern along the major triage categories.
– Interaction between roles increases as severity of the episode increases.
– Different role interaction models were discovered for distinct diagnoses.

Regarding the analysis at the resource level, the main insights are the following:

– A main role-resource interaction pattern was identified where a physician and two nurses work together in most episodes.
– Exceptional behavior for some episodes indicates that also a third nurse and a second physician are sometimes required.
– Prevalence of the main role-resource interaction pattern along the major triage categories.
– Different levels of resource involvement for different diagnoses.

Despite the obtained results when applying the method may seem rather expectable, the proposed method is a contribution by itself given its unbiased replicability at low costs. For instance, the method could be applied before and after a new medical policy is implemented, which allows for analyzing its effects in the collaboration. Also, it can be replicated in different contexts (e.g., diagnoses, triage colors, medical departments, among others) in order to automatically detect strong anomalies with respect to the expected behavior by analysts. However, such analysis should be part of a more robust analysis strategy, involving other techniques and the agents involved. In addition, such analysis could aid on the creation of operational support and decision support systems within the healthcare context [11,33,59].

As future work, there is a plan to enhance the filtering criteria by adding new ER episode attributes, to use additional process mining techniques to identify role interaction models, to implement automatic mechanisms for filtering out less frequent interactions, to analyze independently those less frequent interactions to understand when they happen, and to combine the results obtained in the organizational perspective with results obtained in the control-flow perspective (the

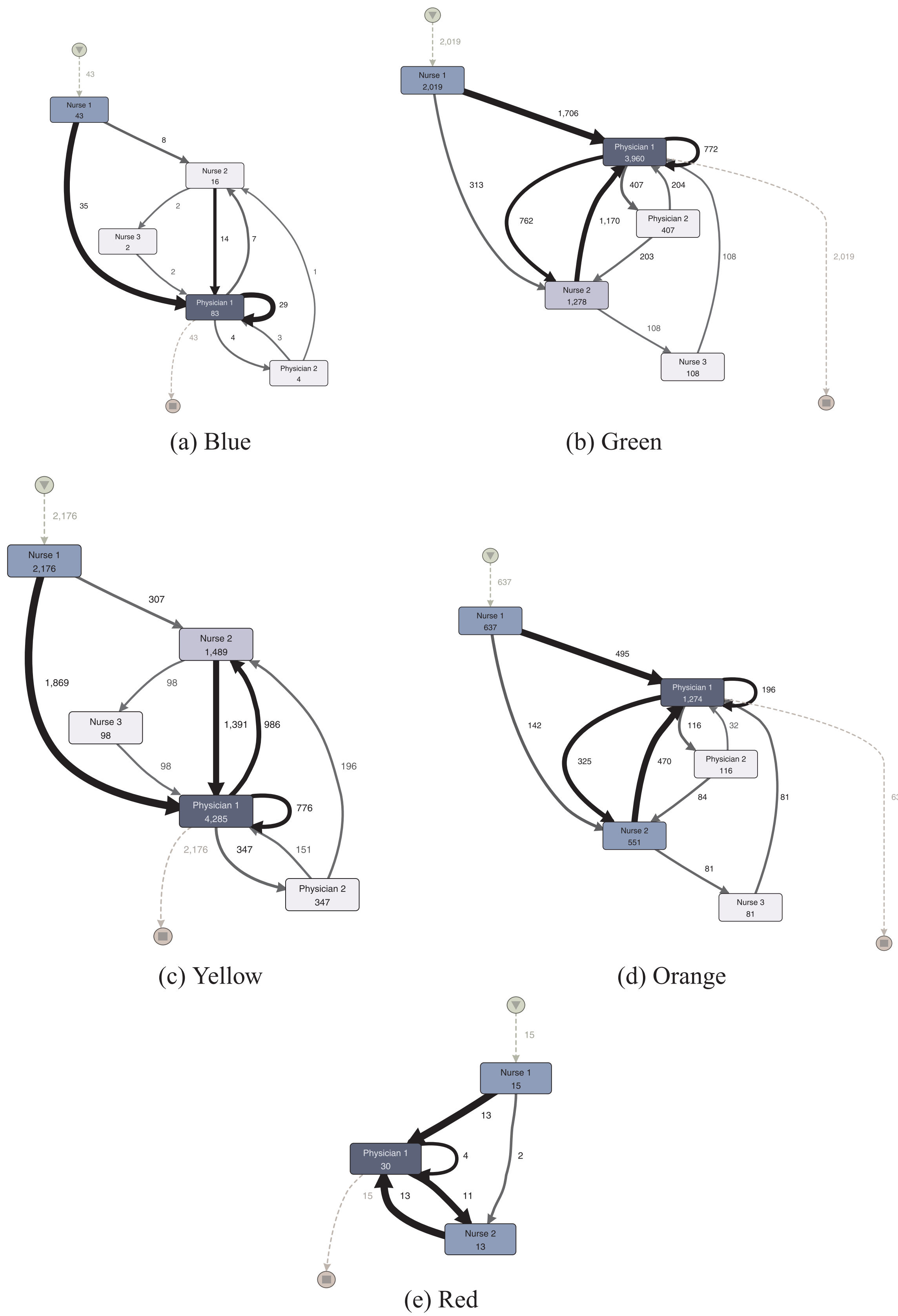
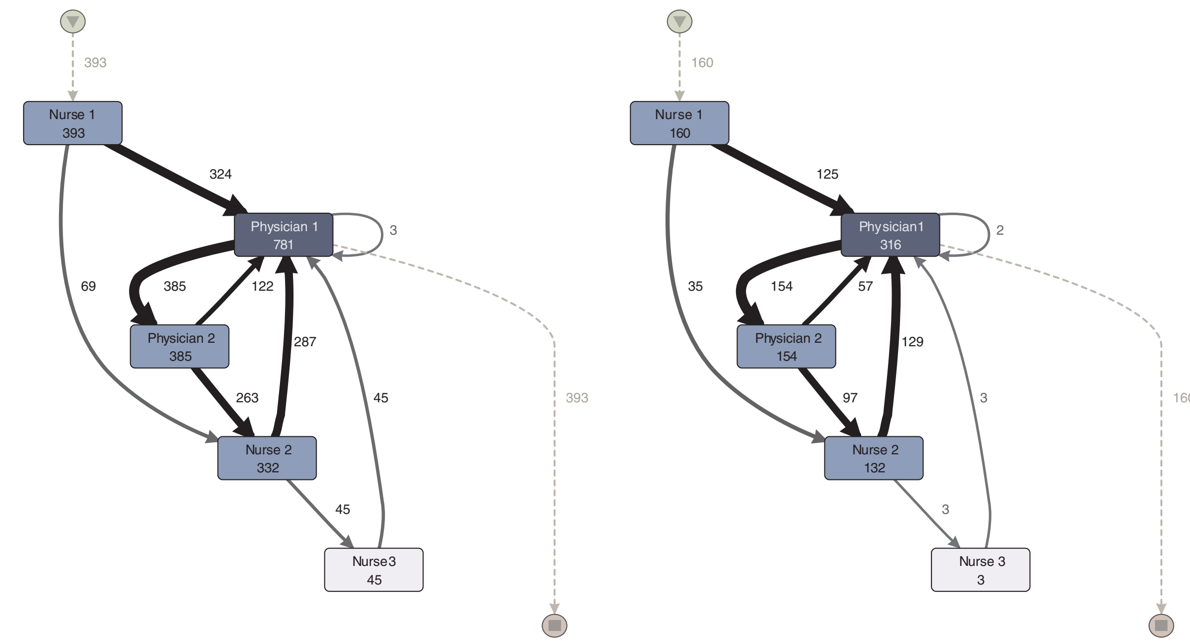


Fig. 13. Role-resource interaction models filtered by triage color.

order in which activities are performed). Since this a first case study, further analysis and validation with other ER experts should also be addressed.

Additional to the previously mentioned analysis, future work may include implementing data mining [60] and text mining [61] techniques to complement the process model perspective with sequential

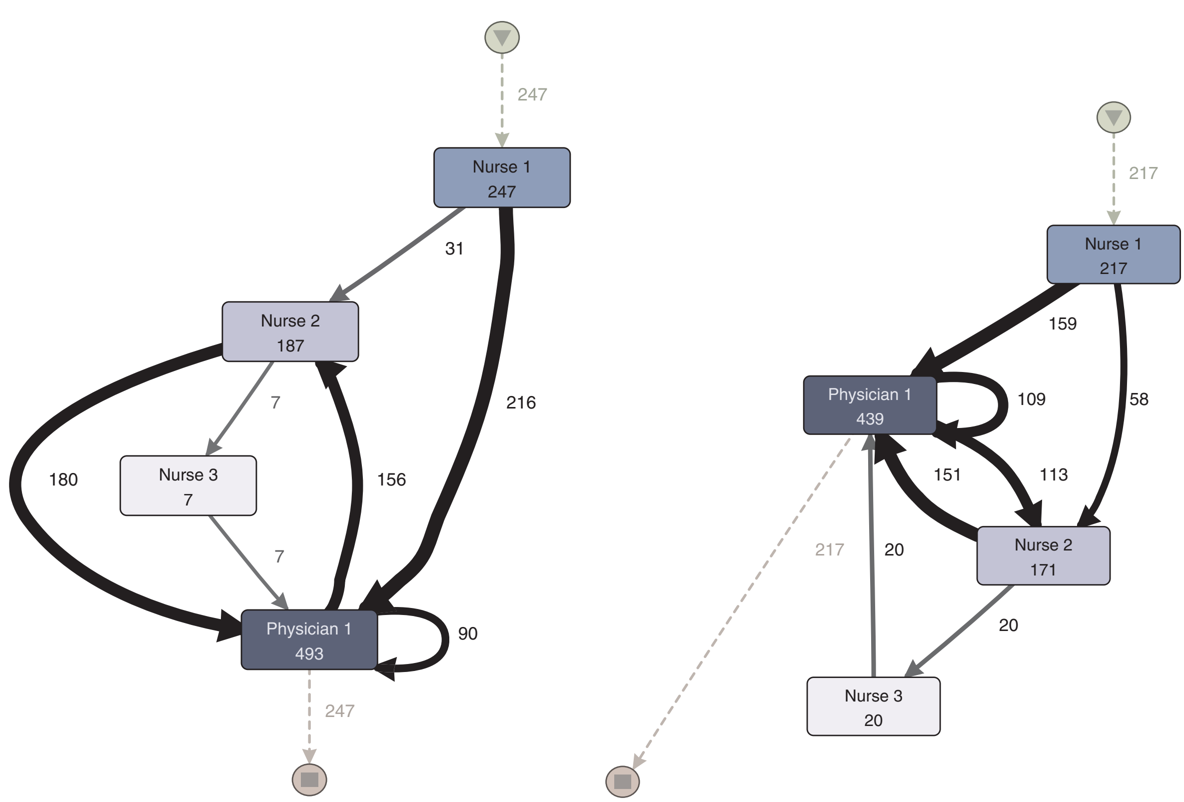


(a) Flu (b) Headache

Fig. 14. Role-resource interaction models in flu and headache episodes.

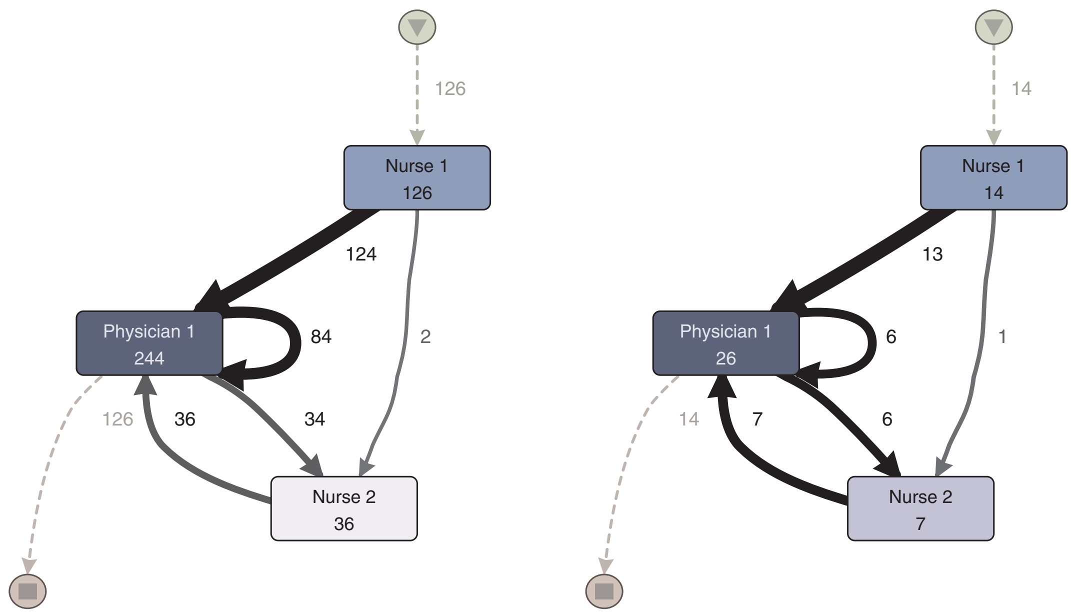
patterns, clustering, data and process trends, and behavior prediction of the extracted data. Also, by combining these different analytical techniques, new knowledge can be provided for more complex and formal Decision Support Systems [62], so as to facilitate organizational process improvement.

Finally, one of the most interesting challenges to address, is how to incorporate healthcare knowledge from the clinical experts into the analysis, enhancing results in real time. This might be implemented using ontologies as information repositories [63] and artificial intelligence techniques as prediction and simulation methods [64].



(a) Abdominal Pain (b) Cold

Fig. 15. Role-resource interaction models in abdominal pain and cold episodes.



(a) Contusions (b) Burns

Fig. 16. Role-resource interaction models in contusions and burns episodes.

Conflict of interests

No conflicts of interest were reported by the authors regarding this study.

Acknowledgments

This project was partially funded by the Ph.D. Scholarship Program of CONICYT – Chile (Doctorado Nacional/2014-63140180 and 2014-63140181), Scholarship Program of CONICIT Costa Rica, Universidad de Costa Rica – Costa Rica and by FONDECYT – Chile Grants 11130577 and 1150365 from the Chilean National Commission on Scientific and Technological Research (CONICYT).

References

[1] D.T. Risser, M.M. Rice, M.L. Salisbury, R. Simon, G.D. Jay, S.D. Berns, M.R. Consortium, et al., The potential for improved teamwork to reduce medical errors in the emergency department, Ann. Emerg. Med. 34 (3) (1999) 373–383.
[2] K.S. Cosby, R. Roberts, L. Palivos, C. Ross, J. Schaider, S. Sherman, I. Nasr, E. Couture, M. Lee, S. Schabowski, et al., Characteristics of patient care management problems identified in emergency department morbidity and mortality investigations during 15 years, Ann. Emerg. Med. 51 (3) (2008) 251–261.
[3] A. Baker, Crossing the quality chasm: a new health system for the 21st century, BMJ: Br. Med. J. 323 (7322) (2001) 1192.
[4] J.C. Morey, R. Simon, G.D. Jay, R.L. Wears, M. Salisbury, K.A. Dukes, S.D. Berns, Error reduction and performance improvement in the emergency department through formal teamwork training: evaluation results of the MedTeams project, Health Serv. Res. 37 (6) (2002) 1553–1581.
[5] K.S. Cosby, A framework for classifying factors that contribute to error in the emergency department, Ann. Emerg. Med. 42 (6) (2003) 815–823.
[6] A. Bryman, Encyclopedia of Social Science Research Methods, 2004.
[7] A.B. Bergman, R.J. Haggerty, The emergency clinic: a study of its role in a teaching hospital, Am. J. Dis. Child. 104 (1) (1962) 36–44.
[8] J. Ryan, Accident and Emergency, Bailey and Love's Short Practice of Surgery 23rd Ed. London, 2000, pp. 270–280.
[9] R. Agarwal, V. Dhar, Big Data, Data Science, and Analytics: The Opportunity and Challenge for is Research, 2014.
[10] W. Raghupathi, V. Raghupathi, Big data analytics in healthcare: promise and potential, Health Inform. Sci. Syst. 2 (1) (2014) 3.
[11] W. van der Aalst, Process Mining: Data Science in Action, Springer, 2016.
[12] R. Mans, W.M.P. van der Aalst, R.J.B. Vanwersch, A.J. Moleman, Process mining in healthcare: data challenges when answering frequently posed questions, in: Process Support and Knowledge Representation in Health Care – BPM 2012 Joint Workshop, ProHealth 2012/KR4HC 2012, Tallinn, Estonia, September 3, 2012, Revised Selected Papers, 2012, pp. 140–153.
[13] E. Rojas, M. Sepúlveda, J. Munoz-Gama, D. Capurro, V. Traver, C. Fernandez-Llatas, Question-driven methodology for analyzing emergency room processes using process mining, Appl. Sci. 7 (3) (2017) 302.
[14] H. Pirnejad, Z. Niazkhani, M. Berg, R. Bal, Intra-organizational communication in healthcare, Methods Inf. Med. 47 (4) (2008) 336–345.
[15] M.T. O'Toole, Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, seventh ed., Miller-Keane, 2003.
[16] M. Dumas, W.M.P. van der Aalst, A.H.M. ter Hofstede, Process-Aware Information Systems: Bridging People and Software Through Process Technology, Wiley, 2005.
[17] R. Mans, W.M.P. van der Aalst, R.J.B. Vanwersch, Process Mining in Healthcare – Evaluating and Exploiting Operational Healthcare Processes, Springer Briefs in Business Process Management, Springer, 2015.
[18] Z. Huang, X. Lu, H. Duan, Resource behavior measure and application in business process management, Expert Syst. Appl. 39 (7) (2012) 6458–6468.
[19] W. Zhao, X. Zhao, Process mining from the organizational perspective, Foundations of Intelligent Systems, Springer, 2014, pp. 701–708.
[20] J. Yan, D. Hu, S.S.Y. Liao, H. Wang, Mining agents' goals in agent-oriented business processes, ACM Trans. Manage. Inf. Syst. 5 (4) (2015) 20:1–20:22.
[21] E. Rojas, J. Munoz-Gama, M. Sepúlveda, D. Capurro, Process mining in healthcare: a literature review, J. Biomed. Inform. 61 (2016) 224–236.
[22] Á. Rebuge, D.R. Ferreira, Business process analysis in healthcare environments: a methodology based on process mining, Inf. Syst. 37 (2) (2012) 99–116.
[23] A. Partington, M. Wynn, S. Suriadi, C. Ouyang, J. Karnon, Process mining for clinical processes: a comparative analysis of four australian hospitals, ACM Trans. Manage. Inform. Syst. (TMIS) 5 (4) (2015) 19.
[24] A. del-Río-Ortega, M. Resinas, C. Cabanillas, A.R. Cortés, Defining and analysing resource-aware process performance indicators, in: Proceedings of the CAiSE'13 Forum at the 25th International Conference on Advanced Information Systems Engineering (CAiSE), Valencia, Spain, June 20th, 2013, 2013, pp. 57–64.
[25] M. Arias, E. Rojas, J. Munoz-Gama, M. Sepúlveda, A framework for recommending resource allocation based on process mining, in: Business Process Management Workshops – BPM 2015, 13th International Workshops, Innsbruck, Austria, August 31–September 3, 2015, Revised Papers, 2015, pp. 458–470.
[26] M. Arias, J. Munoz-Gama, M. Sepúlveda, A multi-criteria approach for team recommendation, in: Business Process Management Workshops – BPM 2016 International Workshops, Rio de Janeiro, Brazil, September 19, 2016, Revised Papers, 2016, pp. 384–396.
[27] W. Van Der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, Springer Science & Business Media, 2011.
[28] N. Russell, W.M.P. van der Aalst, A.H.M. ter Hofstede, D. Edmond, Workflow resource patterns: identification, representation and tool support, in: O. Pastor, J.F. e Cunha (Eds.), CAiSE 2005, LNCS, vol. 3520, Springer, 2005, pp. 216–232.
[29] T. Unger, S. Wagner, Collaboration aspects of human tasks, in: Business Process Management Workshops – BPM 2010 International Workshops and Education Track, Hoboken, NJ, USA, September 13–15, 2010, Revised Selected Papers, 2010, pp. 579–590.
[30] A. Sharma, V. Mansotra, Emerging applications of data mining for healthcare management-a critical review, International Conference on Computing for Sustainable Global Development (INDIACom), 2014, IEEE, 2014, pp. 377–382.
[31] A. Ten Teije, M. Marcos, M. Balser, J. van Croonenborg, C. Duelli, F. van Harmelen, P. Lucas, S. Miksch, W. Reif, K. Rosenbrand, et al., Improving medical protocols by formal methods, Artif. Intell. Med. 36 (3) (2006) 193–209.
[32] D. Buchtela, J. Peleska, M. Zvolsky, J. Zvarova, Medical knowledge representation

system, Stud. Health Technol. Inform. 136 (2008) 377.
[33] L. Karen Marie, From clinical practice guidelines, to clinical guidance in practice–implications for design of computerized guidance, Informat. Technol. Health Care: Socio-Tech. Approach. 2010: Safe Syst. Patient Saf. 157 (2010) 169.
[34] G. Ellingsen, E. Monteiro, A patchwork planet integration and cooperation in hospitals, Comput. Supp. Cooper. Work (CSCW) 12 (1) (2003) 71–95.
[35] M. Berg, W. Schellekens, C. Bergen, Bridging the quality chasm: integrating professional and organizational approaches to quality, Int. J. Qual. Health Care 17 (1) (2005) 75–82.
[36] Y. Engeström, Expansive visibilization of work: an activity-theoretical perspective, Comput. Supp. Cooper. Work (CSCW) 8 (1) (1999) 63–93.
[37] M.S. Granlien, Participation and Evaluation in the Design of Healthcare Work Systems, Roskilde University, Roskilde, 2010.
[38] W.-T. Lin, S.-T. Wang, T.-C. Chiang, Y.-X. Shi, W.-Y. Chen, H.-M. Chen, Abnormal diagnosis of emergency department triage explored with data mining technology: an emergency department at a medical center in taiwan taken as an example, Expert Syst. Appl. 37 (4) (2010) 2733–2741.
[39] A. Ceglowski, L. Churilov, J. Wassertheil, Knowledge discovery through mining emergency department data, Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005. HICSS'05, IEEE, 2005, p. 142c.
[40] H.C. Koh, G. Tan, et al., Data mining applications in healthcare, J. Healthcare Inform. Manage. 19 (2) (2011) 65.
[41] M.E.H. Ong, C.H.L. Ng, K. Goh, N. Liu, Z.X. Koh, N. Shahidah, T.T. Zhang, S. Fook-Chong, Z. Lin, Prediction of cardiac arrest in critically ill patients presenting to the emergency department using a machine learning score incorporating heart rate variability compared with the modified early warning score, Crit. Care 16 (3) (2012) R108.
[42] R.A. Taylor, J.R. Pare, A.K. Venkatesh, H. Mowafi, E.R. Melnick, W. Fleischman, M.K. Hall, Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data–driven, machine learning approach, Acad. Emerg. Med. 23 (3) (2016) 269–278.
[43] M. Laskowski, R.D. McLeod, M.R. Friesen, B.W. Podaima, A.S. Alfa, Models of emergency departments for reducing patient waiting times, PloS One 4 (7) (2009) e6127.
[44] J.-Y. Yeh, W.-S. Lin, Using simulation technique and genetic algorithm to improve the quality care of a hospital emergency department, Expert Syst. Appl. 32 (4) (2007) 1073–1083.
[45] K. Peffers, T. Tuunanen, M.A. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, J. Manage. Inform. Syst. 24 (3) (2007) 45–77.
[46] R. Wirth, J. Hipp, CRISP-DM: towards a standard process model for data mining, in: Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, 2000, pp. 29–39.
[47] W.M.P. van der Aalst, A. Adriansyah, A.K.A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, R.P.J.C. Bose, P. van den Brand, R. Brandtjen, J.C.A.M. Buijs, A. Burattin, J. Carmona, M. Castellanos, J. Claes, J. Cook, N. Costantini, F. Curbera, E. Damiani, M. de Leoni, P. Delias, B.F. van Dongen, M. Dumas, S. Dustdar, D. Fahland, D.R. Ferreira, W. Gaaloul, F. van Geffen, S. Goel, C.W. Günther, A. Guzzo, P. Harmon, A. H.M. ter Hofstede, J. Hoogland, J.E. Ingvaldsen, K. Kato, R. Kuhn, A. Kumar, M.L. Rosa, F.M. Maggi, D. Malerba, R.S. Mans, A. Manuel, M. McCreesh, P. Mello, J. Mendling, M. Montali, H.R.M. Nezhad, M. zur Muehlen, J. Munoz-Gama, L. Pontieri, J. Ribeiro, A. Rozinat, H.S. Pérez, R.S. Pérez, M. Sepúlveda, J. Sinur, P. Soffer, M. Song, A. Sperduti, G. Stilo, C. Stoel, K.D. Swenson, M. Talamo, W. Tan, C. Turner, J. Vanthienen, G. Varvaressos, E. Verbeek, M. Verdonk, R. Vigo, J. Wang, B. Weber, M. Weidlich, T. Weijters, L. Wen, M. Westergaard, M.T. Wynn, Process mining manifesto, in: Business Process Management Workshops – BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I, 2011, pp. 169–194.
[48] R.J.C. Bose, R.S. Mans, W.M. van der Aalst, Wanna improve process mining results? IEEE Symposium on Computational Intelligence and Data Mining (CIDM), 2013, IEEE, 2013, pp. 127–134.
[49] S. Suriadi, R. Andrews, A.H. ter Hofstede, M.T. Wynn, Event log imperfection patterns for process mining: towards a systematic approach to cleaning event logs, Informa. Syst. 64 (2017) 132–150.
[50] C.W. Günther, W.M. Van Der Aalst, Fuzzy mining–adaptive process simplification based on multi-perspective metrics, Business Process Management, Springer, 2007, pp. 328–343.
[51] R.H. Sampieri, C.F. Collado, P.B. Lucio, M.d.l.L.C. Pérez, Metodología de la investigación, McGraw-Hill, 1998.
[52] C.W. Günther, A. Rozinat, Disco: discover your processes, BPM (Demos) 940 (2012) 40–44.
[53] K. Mackway-Jones, C. Robertson, Emergency triage, BMJ-Br. Med. J.-Int. Ed. 314 (7086) (1997) 1056.
[54] I. Mosby, Mosby's Dictionary of Medicine, Nursing & Health Professions, Elsevier Health Sciences, 2013.
[55] E. Rojas, M. Arias, M. Sepúlveda, Clinical processes and its data, what can we do with them, in: International Conference on Health Informatics (HEALTHINF), 2015, pp. 642–647.
[56] W.M.P. van der Aalst, Process-aware information systems: lessons to be learned from process mining, ToPNoC 2 (2009) 1–26.
[57] X. Lu, R. Mans, D. Fahland, W.M.P. van der Aalst, Conformance checking in healthcare based on partially ordered event data, in: Proceedings of the 2014 IEEE Emerging Technology and Factory Automation, ETFA 2014, Barcelona, Spain, September 16–19, 2014, 2014, pp. 1–8.
[58] A. Adriansyah, Aligning Observed and Modeled Behavior (Ph.D. thesis), Eindhoven University of Technology, 2014.
[59] E.S. Berner, Clinical Decision Support Systems vol. 233, Springer, 2007.
[60] I. Yoo, P. Alafaireet, M. Marinov, K. Pena-Hernandez, R. Gopidi, J.-F. Chang, L. Hua, Data mining in healthcare and biomedicine: a survey of the literature, J. Med. Syst. 36 (4) (2012) 2431–2448.
[61] V. Gupta, G.S. Lehal, et al., A survey of text mining techniques and applications, J. Emerg. Technol. Web Intell. 1 (1) (2009) 60–76.
[62] D.J. Power, R. Sharda, Decision support systems, Springer Handbook of Automation, Springer, 2009, pp. 1539–1548.
[63] C.A. Puri, K. Gomadam, P. Jain, P.Z. Yeh, K. Verma, Multiple ontologies in healthcare information technology: motivations and recommendation for ontology mapping and alignment, in: ICBO, 2011.
[64] A. Ramesh, C. Kambhampati, J. Monson, P. Drew, Artificial intelligence in medicine, Ann. Roy. College Surg. Engl. 86 (5) (2004) 334.