Subpopulation Process Comparison and Bottleneck Analysis: A Case Study of Frozen Shoulder

PRIYA NAGUINE, University of Twente, The Netherlands

Frozen shoulder is a condition that limits shoulder movement and causes pain and discomfort. Although there are many treatment options for frozen shoulder, it is not known how these care paths differ among patient groups and whether there are any bottlenecks. This research uses the Process Mining Project Methodology in Healthcare (PM^2HC) and the MIMIC-IV dataset to find the differences in the care paths among different age groups and gender. A seven-step model for bottleneck analysis and resolution was also created. The model has found that there are bottlenecks present in the general procedures followed and medications taken (i.e., continuous infusions and intermittent administrations) during the patients' ICU stays. A limitation of this study is that it has not been evaluated by medical experts in this field so a detailed analysis by them can be a future step.

Additional Key Words and Phrases: adhesive capsulitis, bottleneck analysis, frozen shoulder, healthcare, process mining, subpopulation comparison

1 INTRODUCTION

Characterised by fibrosis and rigidity of the glenohumeral joint, Frozen Shoulder (FS), also known as Adhesive Capsulitis, leads to a loss in the range-of-motion of the shoulder joint [15]. This disease is more prevalent in females than males and is more common from the ages of 40 to 60 [28].

FS develops in three stages: the freezing stage, the frozen stage and the thawing stage [31]. In the **freezing** stage, the patient begins to experience pain, especially at night, as well as restrictions in the range-of-motion of the shoulder joint. The freezing stage is followed by the **frozen** stage in which there is a decrease in the soreness of the joint, but there is a gradual loss in range-of-motion. In the last stage, i.e., the **thawing** stage, there is a further reduction in pain and a slow regain in the range-of-motion.

There are various treatment options for FS, ranging from conservative (e.g., physical therapy) to non-conservative (e.g., capsular release) approaches [26]. The non-conservative alternatives are surgical procedures for which FS patients have to be admitted to the Intensive Care Unit (ICU) in hospitals.

To analyse the procedures followed by FS patients, process mining techniques can be used on the MIMIC-IV database. This database contains data on approximately 300,000 patients that were admitted to a tertiary academic medical centre in Boston, the USA between 2008 and 2019 [18, 20]. Specifically, process mining takes data from hospital information systems (HIS) when applied in the healthcare domain [27]. Event logs are then created using the data from the HIS to show the sequence of processes followed by patients. The event logs created can then be used to find the differences in the care paths followed by subgroups of patients with FS, as well as whether there are bottlenecks in the processes.

This paper is structured as follows. Section 2 will explain the problem statement and the state of the art will be described in section 3. Section 4 will describe the methodology used. Section 5 will answer the research questions and section 6 will look into limitations and future work. Lastly, section 7 concludes this research.

2 PROBLEM STATEMENT

Medical professionals often question whether there is a difference in the treatment procedures followed by subgroups of patients diagnosed with the same disease and where bottlenecks can be found when process mining is applied in the healthcare domain [24]. In this case, a subgroup refers to a group of patients that have a common characteristic, e.g., all female patients diagnosed with FS.

This research aims to answer these questions for patients with FS because this is a matter of concern for medical experts. Since age and gender play a role in the development of FS [21], these will be chosen as the subgroups. There is a limitation that the data set only includes data that was recorded while the patients were in the ICU.

This research paper will answer the following research questions:

(1) What are the differences in the care paths among different groups of frozen shoulder patients?
(2) How can bottlenecks be found and resolved for frozen shoulder patients?

3 STATE OF THE ART

Since this research discovers the care paths followed by FS patients, it is necessary to find out what treatment options are available. In [28], different treatment procedures are proposed for FS patients. The treatment options include oral drugs, intra-articular steroid injections, supra-scapular nerve blocks, hydrodilatation, manipulation under anaesthesia and surgical procedures. However, the literature on FS can be inconsistent due to the lack of awareness on the physiological processes associated with it and how to appropriately distinguish between the phases of the disease [14, 28].

Process mining techniques can be used for various purposes in the healthcare domain, e.g., with BPMN diagrams to get the graph edit distances. [25] used process mining techniques specifically for the process comparison of subgroups. There was a focus on the application of process mining for subpopulation process comparison between patients diagnosed with different types of cancer. In this paper, the subpopulations were compared using graph similarity measures.

The tool BPMNDiffViz [1] can be used to find graph similarity measures. It takes as input two BPMN diagrams and gives the minimal graph edit distance (GED) as a result. The GED can defined as the minimum number of steps required to transform one graph into

[1] https://pais.hse.ru/en/research/projects/CompBPMN/

TScIT 37, July 8, 2022, Enschede, The Netherlands

another [36]. This tool makes use of Business Process Model and Notation (BPMN) 2.0 which is one of the frequently used notations used for process modelling [19]. [35] states that although the use of BPMN diagrams in medicine is a recent development, it can be used to model clinical pathways to teach and train medical staff.

In this research, visual comparison will be used to differentiate between the care paths followed by subgroups of frozen shoulder patients and the tool BPMNDiffViz will be used for that. BPMNDiffViz allows for a choice between six comparison algorithms: Greedy, TabuSearch, Genetic, AStar, Ants and SimulatedAnnealing. [36] compares the algorithms mentioned except Genetic and the conclusion is that the Greedy algorithm gives the best performance results while the TabuSearch algorithm gives more precise and accurate results. The Genetic algorithm only gives an approximation of the GED [32].

4 METHODOLOGY

The methodology to be used in this research is called Process Mining Project Methodology in Healthcare (PM^2HC) [30]. PM^2HC involves 6 phases: planning, extraction, data processing, mining and analysis, evaluation, and improvement and support. An overview of the methodology can be found in figure 1 and the mapping of this research to the six stages of the methodology can be found in sections 4.1 - 4.6.

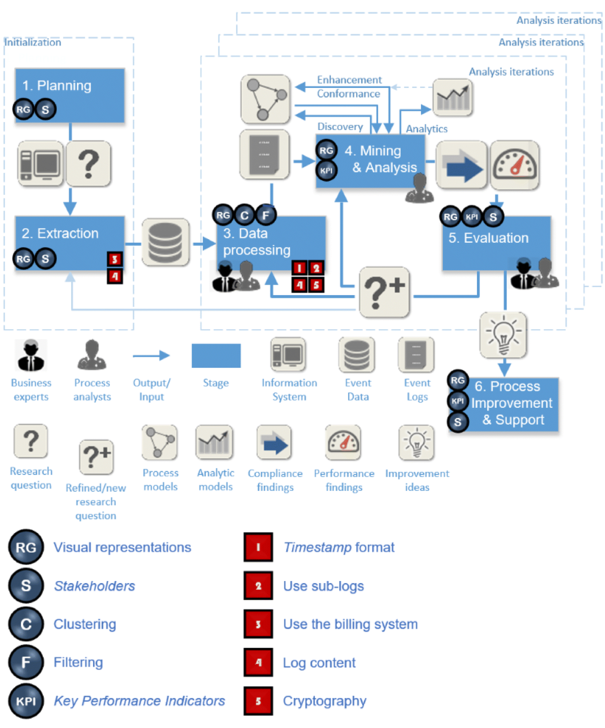

Fig. 1. Process Mining Project Methodology in Healthcare (PM^2HC) [30]

4.1 Planning

In this phase, the subgroups for which the different care paths are investigated were decided, as well as the sequence of events for which bottlenecks are found. Also, background research was performed on frozen shoulder and process mining in healthcare as can be seen in section 3.

4.2 Extraction

In order to get access to the MIMIC-IV database, the CITI program course had to be completed, after which there was a processing time of two weeks for the author's credentials to be verified. Once access to the PhysioNet Clinical databases was granted, the data from the MIMIC-IV database was loaded into Google Cloud Platform BigQuery[2] for the data to be analysed and queries performed on it.

Since the MIMIC-IV database stores the diagnoses given to the patients at the end of their ICU stay using the International Classification of Diseases (ICD) Version 9 and 10 codes, the first step was to find the ICD codes associated with frozen shoulder. This was found in the **D_ICD_DIAGNOSES** table [2] using the keywords *frozen shoulder* and *adhesive capsulitis* for the *long_title*. The ICD codes, versions and their corresponding diagnoses can be found in table 1.

Table 1. ICD codes, versions and diagnoses for frozen shoulder

ICD code	ICD version	Diagnoses
7260	9	Adhesive capsulitis of shoulder
M750	10	Adhesive capsulitis of shoulder
M7500	10	Adhesive capsulitis of unspecified shoulder
M7501	10	Adhesive capsulitis of right shoulder
M7502	10	Adhesive capsulitis of left shoulder

The next step was to find all the patients with the diagnoses given in table 1. This was found in the **DIAGNOSES_ICD** table [1], where the *subject_id* is used to uniquely identify a patient and the *hadm_id* is used to uniquely identify a patient's admission to the hospital. It is important to note that there is a possibility that a patient is given more than one diagnosis associated with frozen shoulder in a single hospitalisation, e.g., a patient can be diagnosed with both M7501 and M7502.

To create individual tables for each subgroup, the **PATIENTS** table [7] was used to get the *anchor_age* and *gender* of the patients. Then, the tables **D_ICD_PROCEDURES** [3] and **PROCEDURES_ICD** [9] were used to get the procedures undergone by the patients in each subgroup. Since a patient can be given more than one diagnosis in a single hospitalisation, the procedures were filtered to include only those used in the diagnosis and treatment of FS. The following keywords, based on the treatment options for FS, were used in the filtering process: *shoulder*, *steroid*, *arthro*, *magnetic resonance imaging*, *rotator*, *physical therapy*, *range of motion testing*, *injection of insulin*.

The tables **D_ICD_DIAGNOSES**, **DIAGNOSES_ICD**, **D_ICD_PROCEDURES** and **PROCEDURES_ICD** were used because they

[2] https://cloud.google.com/bigquery

contain the diagnoses given to the patients and the procedures that the patients go through. This information is used by the hospital to bill the patients for the care provided and PM^2HC recommends using this data as it is more reliable [30].

To find the bottlenecks in the medications taken (i.e., continuous infusions and intermittent administrations) and the general procedures undergone by patients during their ICU stay, the tables **INPUTEVENTS** [6] and **PROCEDUREEVENTS** [8] were used, respectively. These tables contained the *subject_id*, *hadm_id*, *starttime*, *endtime* and *ordercategoryname*. The *ordercategoryname* describes the medication taken and procedure followed, respectively.

To apply process mining algorithms to the data, the cases, events, start times and end times have to be defined. For both the subgroup process comparison and bottleneck analysis, a case is a patient's admission to the hospital and the events are the procedures that the patients were billed for. Since the start and end times were not stored for the subgroup process comparison, the sequence number was used instead to indicate the order in which the procedures were carried out.

Once the required data was stored in the appropriate BigQuery tables, the tables were exported as CSV files for later usage.

4.3 Data processing

In this phase, the CSV files on the subgroups were entered into ProM, converted into XES files and visualised using the **LogVisualiser (LogDialog)** plugin. Table 2 gives an overview on the number of cases and events per subgroup, given by the LogDialog. Also, further filtering was required to find the differences in care paths between the different patient groups. This was done using the **Filter Log on Event Attribute Values** plugin, where certain procedures were filtered out from the care paths.

Table 2. Number of cases and events per subgroup

Subgroup	#Cases	#Events
Female *	29	61
Male *	34	55
Age below 40 **	8	18
Age between 40 and 60 **	39	73
Age above 60 **	16	25

* Includes FS patients from all age groups
** Includes FS patients from both genders

4.4 Mining and Analysis

This phase involved finding the differences in the care paths between the different subgroups and the bottlenecks in the medications taken and the procedures followed by patients during their ICU stays. To do this, process models were created in ProM [3] and Disco [4].

The process models for the subgroup process comparison were created using the following plugins in ProM: **Mine with Inductive visual Miner**, **Mine Petri net with Inductive Miner** and **Convert Petri net to BPMN diagram**.

[3] https://www.promtools.org/doku.php
[4] https://fluxicon.com/disco/

The Inductive Miner was chosen because it gives the best fitness, i.e., the degree by which the process models generated can recreate the cases in the event log [12]. Firstly, the plugin **Mine with Inductive visual Miner** was used because it can create animations showing the order in which the processes occur; it was used with the *activities* slider set to 1 and the *paths* slider set to 0.8. These settings were chosen so that the Petri net and the Inductive visual Miner models are equivalent. Secondly, **Mine Petri net with Inductive Miner** was used to create static process models which can be used for visual comparison, with a *noise threshold* of 0.2 to allow for small deviations. Lastly, in order to convert the Petri net models into BPMN diagrams so that they can be loaded into BPMNDiffViz to get the GED, **Convert Petri net to BPMN diagram** was used.

The **Performance** function was used in Disco for the bottleneck analysis of the procedures followed and medications taken. The first metric chosen was the **median duration** and the second metric was the **absolute frequency**. Since there is a large variation in the duration of the cases, the median duration was chosen as it is less likely to be affected by outliers [34]. The *activities* slider was set to 100% and *paths* slider was set to 80% to filter out small deviations. The **Attribute** filter was also used so that individual cases could be investigated.

Details on the results found and their interpretation can be found in section 5.

4.5 Evaluation

In this phase, the insights obtained in the previous phase were used to suggest improvements to the system by considering the bottlenecks. Further details on this phase can be found in section 6.

4.6 Improvement and support

In this phase, the stakeholders, e.g., medical professionals, decide on the path to be followed to implement the improvements. This step was not performed as this research was not performed in partnership with a treatment provider.

5 FINDINGS

The process models created in ProM and Disco for the subgroup process comparison and bottleneck analysis can be found in the author's GitHub repository [5].

5.1 Research Question 1

To answer research question 1, two subquestions will be formulated:

(1) What are the differences in the care paths followed by male and female frozen shoulder patients?
(2) What are the differences in the care paths followed by frozen shoulder patients aged below 40, between 40 and 60 inclusive, and above 60?

When comparing the care paths of the subgroups, three keywords will be used. Firstly, **parallel** will be used when two procedures occur but the order in which they occur does not matter. Secondly, **sequence** is used when one procedure follows another. Lastly, **exclusive** will be used when only one of two procedures can occur.

[5] https://github.com/PriyaNaguine/Complete-Process-Models-Frozen-Shoulder

Also, visual comparison is performed in BPMNDiffViz using the **TabuSearch** algorithm with *maximum expansions* and *tabu list size* set to 100 as this gives precise results faster than other algorithms [36]. The activities in the BPMN diagrams are encoded with different colours: blue denotes elements that match between the subgroups, green denotes elements that should be added to transform one diagram into the other and red denotes elements that should be deleted to transform one diagram into the other.

5.1.1 Male vs Female. Visual comparison was done in BPMNDiffViz for the care paths followed by male and female FS patients, resulting in a final score of 167 using the TabuSearch algorithm. The statistics can be found in table 3. Also, table 4 shows the procedures that are only performed on either female or male FS patients, but not both.

Table 3. Statistics for the comparison of the care paths between male and female patients

	Percentage of elements	**Number of elements**
Matched elements	37%	35
Deleted elements *	33%	31
Added elements *	30%	28

* Refer to table 4 for the differences in elements

Table 4. Procedures performed on either male or female FS patients

Procedure	**Female**	**Male**
Drainage of Right Shoulder Joint, Percutaneous Approach, Diagnostic		✓
Excision of Left Shoulder Bursa and Ligament, Percutaneous Endoscopic Approach	✓	
Excision of Right Shoulder Joint, Percutaneous Endoscopic Approach		✓
Other total shoulder replacement	✓	
Release Right Shoulder Joint, External Approach		✓
Repair of recurrent dislocation of shoulder	✓	
Repair Right Shoulder Joint, Percutaneous Endoscopic Approach		✓
Repair Right Shoulder Tendon, Open Approach		✓

The procedure **"Release Right Shoulder Joint, Open Approach"** is performed on both male and female FS patients. However, when it is performed on male patients, it follows "Repair Right Shoulder Tendon, Open Approach". In female patients, it takes place after "Replacement of Right Shoulder Joint with Reverse Ball and Socket Synthetic Substitute, Open Approach".

The procedure **"Rotator cuff repair"**, if performed on male patients, it is always the first procedure. Relating it to the procedure "Other local excision or destruction of lesion of joint, shoulder", the procedures are exclusive for male patients. For female patients, these two procedures can be done in sequence.

The procedure **"Other arthrotomy, shoulder"**, if performed, is always the first procedure for male patients. When relating this to "Other repair of shoulder", it can be performed in sequence for both genders, but on female patients, "Other arthrotomy, shoulder" takes place after "Other repair of shoulder".

The procedure "Skeletal x-ray of shoulder and upper arm" is only performed on male patients while "Magnetic resonance imaging of other and unspecified sites" is only performed on female patients. These procedures are not performed in combination with other procedures.

The procedure **"Other repair of shoulder"** can be done in parallel with "Division of joint capsule, ligament, or cartilage, shoulder" in male patients while in female patients, these procedures are performed in sequence. Furthermore, it is performed in sequence with "Rotator cuff repair" in male FS patients. However, in female patients, these processes are exclusive. This can be seen in figures 2 and 3.

As can be seen in figures 2 and 3, the procedure **"Synovectomy, shoulder"** is always the last process in male FS patients, in case it is performed. In female patients, it is exclusive with "Rotator cuff repair", while in male patients, they can occur in sequence, where "Rotator cuff repair" is the first procedure and "Synovectomy, shoulder" is the last procedure to take place.

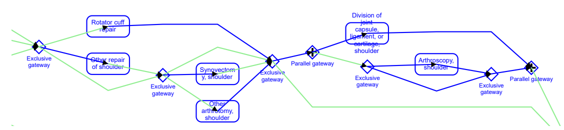

Fig. 2. Snapshot of the BPMN diagram for female FS patients

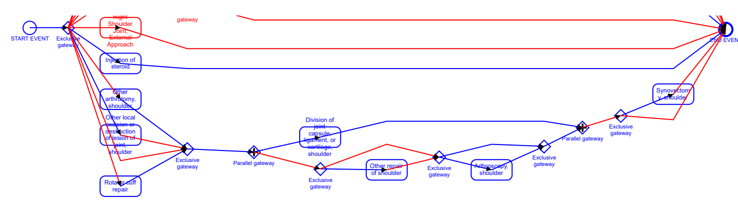

Fig. 3. Snapshot of the BPMN diagram for male FS patients

5.1.2 Age below 40 vs Age between 40 and 60. Visual comparison was done in BPMNDiffViz for the care paths followed by patients aged below 40 and patients aged between 40 and 60, resulting in a final score of 135 using the TabuSearch algorithm. The statistics can be found in table 5. Also, table 6 shows the procedures that are only performed on either patients aged below 40 or patients aged between 40 and 60, but not both.

For the procedure **"Release Shoulder Joint"**, patients aged below 40 follow a "percutaneous endoscopic approach" while those aged between 40 and 60 follow an "external approach".

Table 5. Statistics for the comparison of the care paths between patients aged below 40 and patients aged between 40 and 60

	Percentage of elements	**Number of elements**
Matched elements	49%	35
Deleted elements *	10%	7
Added elements *	41%	29

* Refer to table 6 for the differences in elements

Table 6. Procedures performed on either patients aged below 40 or patients aged between 40 and 60

Procedure	**Age below 40**	**Age between 40 and 60**
Drainage of Right Shoulder Joint, Percutaneous Approach, Diagnostic		✓
Excision of Left Shoulder Bursa and Ligament, Percutaneous Endoscopic Approach	✓	
Excision of Right Shoulder Joint, Percutaneous Endoscopic Approach		✓
Magnetic resonance imaging of other and unspecified sites		✓
Other total shoulder replacement		✓
Repair of recurrent dislocation of shoulder		✓
Repair Right Shoulder Tendon, Open Approach		✓
Rotator cuff repair		✓

As can be seen in figures 4 and 5, the procedure **"Other arthrotomy, shoulder"** is performed in sequence with "Other repair of shoulder" for patients below the age of 40 while for patients aged between 40 and 60, the procedures are exclusive. Relating it to "Synovectomy, shoulder", they are carried out in sequence for the age group between 40 and 60 but they are exclusive for patients aged below 40.

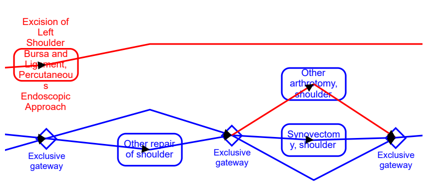

Fig. 4. Snapshot of the BPMN diagram for FS patients aged below 40

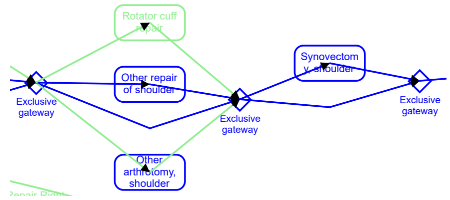

Fig. 5. Snapshot of the BPMN diagram for FS patients aged between 40 and 60

5.1.3 *Age between 40 and 60 vs Age above 60.*

Visual comparison was done in BPMNDiffViz for the care paths followed by patients aged above 60 and patients aged between 40 and 60, resulting in a final score of 166 using the TabuSearch algorithm. The statistics can be found in table 7. Also, table 8 shows the procedures that are only performed on either patients aged above 60 or patients aged between 40 and 60, but not both.

Table 7. Statistics for the comparison of the care paths between patients aged above 60 and patients aged between 40 and 60

	Percentage of elements	**Number of elements**
Matched elements	34%	30
Deleted elements *	27%	24
Added elements *	39%	34

* Refer to table 8 for the differences in elements

The procedure **"Release Right Shoulder Joint"** follows an "open approach" in patients aged above 60 while an "external approach" is followed by patients aged between 40 and 60.

Patients aged above 60 carry out "Skeletal x-ray of shoulder and upper arm" while those aged between 40 and 60 conduct "Magnetic resonance imaging of other and unspecified sites". As was the case with the gender subgroups, they are not performed in combination with other procedures.

The procedure **"Division of joint capsule, ligament, or cartilage, shoulder"** is optional only for those patients aged above 60. It can be done in parallel with "Synovectomy, shoulder" in patients aged above 60 while they are carried out in sequence in patients aged between 40 and 60. This can be seen in figures 6 and 7.

As can be seen in figures 6 and 7, the procedure **"Synovectomy, shoulder"** is performed in sequence with "Rotator cuff repair" in patients aged between 40 and 60 while they are exclusive in patients aged above 60. The same holds for "Other repair of shoulder" and "Arthroscopy, shoulder" when related to "Synovectomy, shoulder".

5.2 Research Question 2

To answer research question 2, two subquestions will be formulated:

(1) What are the bottlenecks in the medications taken (i.e., continuous infusions and intermittent administrations) by frozen shoulder patients during their ICU stay?
(2) What are the bottlenecks in the procedures followed by frozen shoulder patients during their ICU stay?

Table 8. Procedures performed on either patients aged above 60 or patients aged between 40 and 60

Procedure	Age between 40 and 60	Age above 60
Drainage of Right Shoulder Joint, Percutaneous Approach, Diagnostic	✓	
Excision of Right Shoulder Joint, Percutaneous Endoscopic Approach	✓	
Injection of steroid		✓
Magnetic resonance imaging of other and unspecified sites	✓	
Other arthrotomy, shoulder	✓	
Other total shoulder replacement	✓	
Repair of recurrent dislocation of shoulder	✓	
Repair Right Shoulder Joint, Percutaneous Endoscopic Approach	✓	
Replacement of Right Shoulder Joint with Reverse Ball and Socket Synthetic Substitute, Open Approach		✓
Skeletal x-ray of shoulder and upper arm		✓

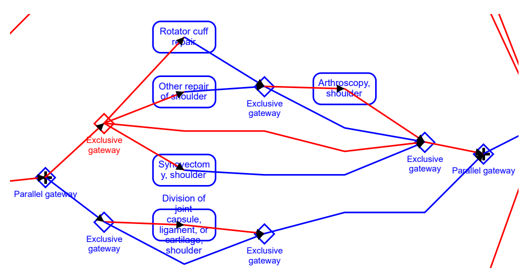

Fig. 6. Snapshot of the BPMN diagram for FS patients aged above 60

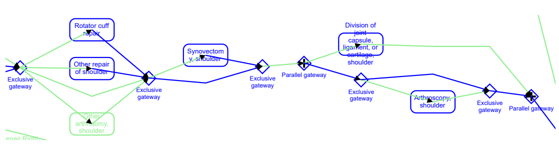

Fig. 7. Snapshot of the BPMN diagram for FS patients aged between 40 and 60

Table 9 shows the statistics, given by Disco, on the data for the procedures followed and medications taken by FS patients.

For the bottleneck analysis of the procedures followed and medications taken, it is necessary to identify the bottlenecks and give

Table 9. Statistics for the bottleneck analysis of the procedures followed and medications taken by FS patients during their ICU stays

	Procedures followed	Medications taken
Events	639	7392
Cases	39	39
Activities	10	15
Median case duration	50.9 hours	48.5 hours
Mean case duration	5.4 days	5.3 days

recommendations on how to resolve them. A few studies on bottleneck analysis in healthcare will be analysed and a general model on bottleneck analysis and resolution will be created based on that.

Firstly, there is a study that was carried out in collaboration with 12 hospitals in China with the purpose of identifying the bottlenecks in implementing basic precautions in the medical centres [23]. In this study, the bottlenecks were identified, along with the reasons behind it, and recommendations were given on how to resolve each bottleneck.

Secondly, participatory action research was used in [16] to find the bottlenecks in healthcare systems in Nigeria. A four-step route was used for bottleneck analysis and performance improvement. This included identifying the bottlenecks, discovering its root causes, finding solutions and strategies to combat the bottlenecks and lastly, consulting the relevant stakeholders to validate the findings. However, it does mention that finding the bottlenecks and solutions to resolve them does not ensure optimal performance.

Lastly, a simulation study was conducted in the University of Kentucky Chandler Hospital to find the most favourable assignment of workforce and resources for improvement [13]. In this study, the main bottleneck was found and there were recommendations made on the number of staff members required as well as the number of machines.

In these studies, it was possible to give recommendations on how to resolve the bottlenecks as the amount of resources that were being used was known. Since this research is not conducted in collaboration with a treatment provider, it is not possible to give concrete suggestions on the amount of resources required, but a model for bottleneck analysis can be constructed based on these studies. The seven-step model created can be found in figure 8. This model first finds the bottlenecks, after which their causes, impact and dependencies are found. Then, the resolution mechanisms are proposed and discussed with the relevant stakeholders before the resolution process is started.

It should be mentioned that in order to answer the research questions, the bottleneck analysis and resolution will be done solely in the context of frozen shoulder, by considering the seven-step model.

5.2.1 Bottlenecks: Medications taken. The largest bottleneck is **13-Enteral Nutrition** with a duration of 7 hours. There is a waiting time of 23.5 hours between **13-Enteral Nutrition** and **16-Pre Admission/Non-ICU**.

The procedure **Enteral Nutrition** involves providing nutrients to the patients directly into the stomach using tubes which can be

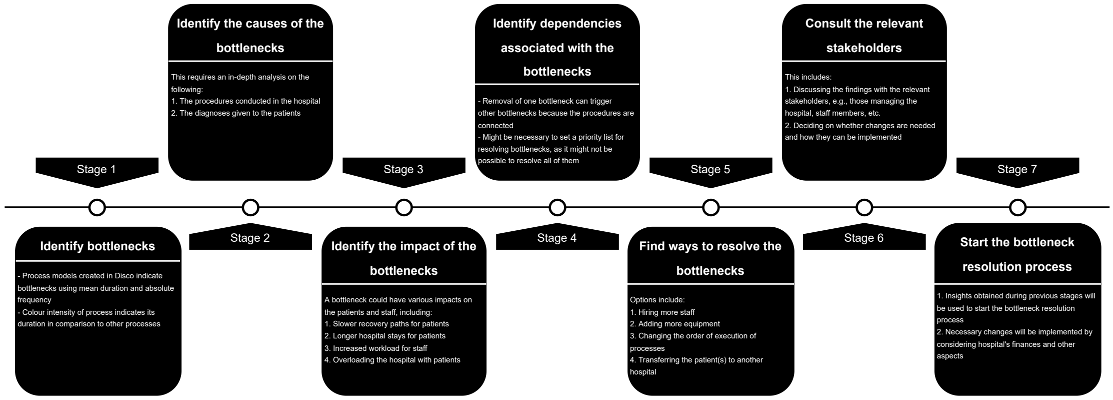

Fig. 8. Bottleneck analysis and resolution model

inserted either through the nose or through a small incision made in the abdomen [10]. This procedure is used under circumstances in which the patient has difficulty swallowing, neurological problems, gastrointestinal problems or injuries in the digestive track. Since this procedure is not relevant to FS, it is a bottleneck due to another diagnosis given to the patient. As a result, it can be disregarded as a bottleneck for FS patients.

The medications relevant to FS are **07-Blood Products**, **09-Antibiotics (Non IV)** and **11-Prophylaxis (Non IV)**. **Prophylaxis** is defined as "an attempt to prevent disease" and **antibiotics** is defined as "a drug used to treat infections caused by bacteria and other microorganisms" by NCI's Dictionary of Cancer Terms [6]. Blood product is defined as "any therapeutic substance derived from human blood, including whole blood and other blood components for transfusion, and plasma-derived medicinal products" by the World Health Organisation [7].

Since platelet-rich plasma (PRP) is a treatment option for other orthopedic diseases, it could be used for FS too as suggested by the experimental study in [17]. The study concluded that PRP injections lowered the severity level of some parts of FS without side effects. Future work included investigating PRP injections as a potential prophylaxis of FS in shoulders that are either immobilised or have undergone surgery, or as a treatment for FS in the early stages. Moreover, a clinical trial made use of intra-articular Doxycycline, an antibiotic which was injected once biweekly for 4 weeks, as a treatment for FS [22].

Given the data from the MIMIC-IV database and the information gathered on FS, only the processes **07-Blood Products**, **09-Antibiotics (Non IV)** and **11-Prophylaxis (Non IV)** are relevant to FS and have the potential of becoming bottlenecks. This is in alignment with the process models as the arrows going into **09-Antibiotics (Non IV)** and **11-Prophylaxis (Non IV)** are thicker than others and **11-Prophylaxis (Non IV)** has a high frequency of 176. Because these processes only take 60 seconds, there is a low chance of them becoming bottlenecks. However, **07-Blood Products** has a duration of 60 minutes and a frequency of 43, increasing its chances of being a bottleneck.

5.2.2 Bottlenecks: Procedures followed. The biggest bottleneck is **Invasive Lines** with a duration of 44 hours. There is a waiting time of 29.4 days between **Intubation/Extubation** and **Invasive Lines**.

The procedure **Invasive Lines** involves inserting medical devices into the patient's body, e.g., urinary catheters, intravenous lines, PICC lines and central lines [11]. Intravenous lines through the patient's arm can be used to deliver anesthesia, with the possibility of inserting a tube into the patient's mouth, down the trachea, to ensure that the patient gets enough oxygen during the surgical procedure [5]. This makes **Intubation/Extubation** a procedure relevant to FS patients.

A treatment option for FS that makes use of this is manipulation under anaesthesia. In this procedure, the surgeon manipulates the shoulder while the patient is anaesthetised, with the aim of increasing the range-of-motion of the shoulder [29].

Another procedure that is relevant to FS patients is **Imaging**. The process models on the care paths of the different patient groups include the processes "Skeletal x-ray of shoulder and upper arm" and "Magnetic resonance imaging of other and unspecified sites". These procedures, if performed, are always the first and only procedures to be carried out. This could mean that imaging procedures are carried out on FS patients to get an initial diagnosis and possibly rule out or confirm shoulder problems [4].

Therefore, there are three procedures that are relevant to frozen shoulder: **Imaging**, **Intubation/Extubation** and **Invasive Lines**. From these procedures, the greatest bottleneck is **Invasive Lines** and it cannot be eliminated as it is necessary when manipulation under anaesthesia is performed. The procedures **Imaging** and **Intubation/Extubation** have a duration of 60 seconds so there is a low probability of them becoming bottlenecks.

6 DISCUSSION

The treatment procedures for FS can be analysed through patient evaluation, public sources evaluation and treatment provider evaluation.

[6] https://www.cancer.gov/publications/dictionaries/cancer-terms
[7] https://www.who.int/health-topics

When performing background research on frozen shoulder and process mining, there was limited information about it in scientific papers. Web pages gave a clearer idea on how frozen shoulder is diagnosed and what is implemented in practice for its treatment. This can be explained by the gap between medical research and what is done in practice, as it can take up to 17 years for clinical discoveries to be applied in practice [33]. Practice-based research can be used to find the causes of this gap as this is carried out by doctors and takes place in the room in which the patients have their doctor's visit [37]. Carrying out practice-based research could mean that treatment procedures and medications are tried out with patients before research papers are written on it, explaining why there could be more information about FS on web pages and blog posts as compared to scientific papers.

Furthermore, bottleneck resolution could not be done in terms of the exact resources (i.e., staff and equipment) needed as this research was not conducted in collaboration with a treatment provider. However, general suggestions can be given for bottleneck resolution. In terms of resources, it might be necessary to increase the number of staff in the hospital, as lack of staff could lead to long waiting times, slower execution of procedures and increased workload for the current staff. With regards to equipment, specifically medical imaging equipment and hospital beds, an increase in number could be essential as lack of such equipment could lead to long queues because patients cannot be admitted into the hospital without a sufficient amount of beds. However, if the hospital is not financially stable, it is necessary to look for solutions that do not exceed the hospital's budget. This can include changing the order of the processes in the patients' care paths, but whether this can be done or not depends on the patients' health. In the worst case, it could be necessary to transfer the patients to another hospital, because long waiting times can be detrimental to the patients. Again, since the research is not done with a hospital, it is not feasible to test how changing the order of the processes or transferring the patients would affect the overall duration.

6.1 Limitations and Future Work

The dataset only includes data on the patients while they are in the ICU and only on patients admitted to one hospital. Because of this, there can be bias towards the practices conducted at the particular medical institute. These factors can impact the generalisability of the findings in this research to all FS patients. In the future, this research could be expanded by considering data from multiple treatment providers, which will allow for data on more patients. Also, other subgroups of FS patients could be explored and the differences among their care paths.

Another limitation is that it cannot be ensured that the procedures followed and medications taken were used solely in the treatment of FS. Considering the manner in which the MIMIC-IV database is structured, it is not possible to distinguish between the procedures followed and medication taken for the different diagnoses in order to perform bottleneck analysis. For the subgroup process comparison, differentiating between the procedures performed for different diagnoses was possible due to the keywords gathered during the background research phase but there is no guarantee that the list of keywords is exhaustive.

Furthermore, since there are over 100,000 unique diagnoses in the MIMIC-IV dataset, it is possible for this research to be done on another disease. If subgroups can be defined for another disease, it is possible to use process mining techniques to find the differences in the care paths among them. Also, since timestamps are included for the procedures followed and medications taken by the patients, it should be possible to also do bottleneck analysis for patients diagnosed with another diseases, taking into account that it is not possible to distinguish between the diseases for bottleneck analysis.

Moreover, since this research was not done in collaboration with a treatment provider, the last step of the PM^2HC was not performed. In this direction, future work could include expanding on this research and performing it in partnership with treatment providers so that phase *improvement and support* can be performed.

7 CONCLUSION

This research applies process mining techniques on data from the MIMIC-IV dataset, specifically on patients diagnosed with frozen shoulder. In particular, the differences in the care paths followed by male and female FS patients were found, as well as the differences in the care paths followed by patients aged below 40, between 40 and 60, and above 60. Also, a seven-step framework was created for bottleneck analysis and resolution. In the context of FS, bottleneck analysis was performed on the procedures followed and medications taken (i.e., continuous infusions and intermittent administrations) by the patients.

For the subpopulation process comparison, the highest graph edit distance was 167 for the care paths of male and female FS patients, followed by 166 for the care paths of patients aged above 60 and between 40 and 60. Finally, the GED for the care paths of patients aged below 40 and between 40 and 60 was 135.

There were bottlenecks present in the medications taken and procedures followed by FS patients. The biggest bottleneck in the medications taken was **13-Enteral Nutrition** but this is not relevant to FS and can therefore be eliminated. The biggest bottleneck in the procedures followed was **Invasive Lines** which could not be eliminated as it is relevant when FS patients undergo manipulation under anaesthesia.

By considering the research questions, future work could include conducting the research in collaboration with treatment providers, conducting subgroup process comparison and bottleneck analysis for other diseases and conducting the research with data from different medical institutes to reduce bias and improve generalisability.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Faiza A. Bukhsh, for guiding me through this research project, providing valuable input and always encouraging me to do my best.

REFERENCES

[1] 2020. diagnoses_icd. https://mimic.mit.edu/docs/iv/modules/hosp/diagnoses_icd/ [Online; accessed 17. May. 2022].
[2] 2020. d_icd_diagnoses. https://mimic.mit.edu/docs/iv/modules/hosp/d_icd_diagnoses/ [Online; accessed 17. May. 2022].

[3] 2020. d_icd_procedures. https://mimic.mit.edu/docs/iv/modules/hosp/d_icd_procedures/ [Online; accessed 17. May. 2022].

[4] 2020. Frozen Shoulder. https://www.mayoclinic.org/diseases-conditions/frozen-shoulder/diagnosis-treatment/drc-20372690 [Online; accessed 03. Jun. 2022].

[5] 2020. General anesthesia. https://www.mayoclinic.org/tests-procedures/anesthesia/about/pac-20384568 [Online; accessed 03. Jun. 2022].

[6] 2020. Inputevents. https://mimic.mit.edu/docs/iv/modules/icu/inputevents/ [Online; accessed 17. May. 2022].

[7] 2020. Patients table. https://mimic.mit.edu/docs/iv/modules/core/patients/ [Online; accessed 17. May. 2022].

[8] 2020. procedureevents. https://mimic.mit.edu/docs/iv/modules/icu/procedureevents/ [Online; accessed 17. May. 2022].

[9] 2020. procedures_icd. https://mimic.mit.edu/docs/iv/modules/hosp/procedures_icd/ [Online; accessed 17. May. 2022].

[10] 2021. Home Enteral Nutrition. https://www.mayoclinic.org/tests-procedures/home-enteral-nutrition/about/pac-20384955 [Online; accessed 02. Jun. 2022].

[11] 2022. Invasive devices. https://www.sepsis.org/sepsisand/invasive-devices/ [Online; accessed 19. May. 2022].

[12] Alejandro Bogarín, Rebeca Cerezo, and Cristóbal Romero. 2018. Discovering learning processes using Inductive Miner: A case study with Learning Management Systems (LMSs). *Psicothema* 30 (Aug. 2018), 322–329. https://doi.org/10.7334/psicothema2018.116

[13] Stuart Brenner, Zhen Zeng, Yang Liu, Junwen Wang, Jingshan Li, and Patricia Howard. 2010. Modeling and Analysis of the Emergency Department at University of Kentucky Chandler Hospital Using Simulations. *Journal of emergency nursing: JEN : official publication of the Emergency Department Nurses Association* 36, 4 (July 2010), 303–310. https://doi.org/10.1016/j.jen.2009.07.018

[14] Daniel de la Serna, Santiago Navarro-Ledesma, Fany Alayón, Elena López, and Leo Pruimboom. 2021. A Comprehensive View of Frozen Shoulder: A Mystery Syndrome. *Frontiers in Medicine* 8 (2021). https://doi.org/10.3389/fmed.2021.663703

[15] Giovanni Maria D'Orsi, Alessio Giai Via, Antonio Frizziero, and Francesco Oliva. 2012. Treatment of adhesive capsulitis: a review. *Muscles, ligaments and tendons journal* 2, 2 (2012), 70–78.

[16] Ejemai Eboreime, Obinna Idika, Kasarachi Omitiran, Lauretta Eboreime, and Latifat Ibisomi. 2019. Primary healthcare planning, bottleneck analysis and performance improvement: An evaluation of processes and outcomes in a Nigerian context. *Evaluation and Program Planning* 77 (Sept. 2019). https://doi.org/10.1016/j.evalprogplan.2019.101712

[17] Oscar Feusi, Agnieszka Karol, Thea Fleischmann, Brigitte von Rechenberg, Samy Bouaicha, Clément Werner, and Thorsten Jentzsch. 2022. Platelet-rich plasma as a potential prophylactic measure against frozen shoulder in an in vivo shoulder contracture model. *Archives of Orthopaedic and Trauma Surgery* 142 (March 2022). https://doi.org/10.1007/s00402-020-03617-x

[18] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P.C. Ivanov, R. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, and H.E. Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation [Online]* 101, 23 (2000), e215–e220.

[19] S.Y. Ivanov, Anna Kalenkova, and Wil Aalst. 2015. BPMNDiffViz: A tool for BPMN Models comparison. 1418 (Jan. 2015), 35–39.

[20] Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. 2021. MIMIC-IV (version 1.0). https://doi.org/10.13026/s6n6-xd98

[21] Rinco Koorevaar, Esther Riet, Marcel Ipskamp, and Sjoerd Bulstra. 2017. Incidence and prognostic factors for postoperative frozen shoulder after shoulder surgery: a prospective cohort study. *Archives of Orthopaedic and Trauma Surgery* 137 (March 2017). https://doi.org/10.1007/s00402-016-2589-3

[22] John E Kuhn. 2022. Intra-articular Doxycycline: A Novel Treatment of Adhesive Capsulitis (DOXY). https://clinicaltrials.gov/ct2/show/study/NCT03479502 [Online; accessed 03. Jun. 2022].

[23] Chunqing Lin, Li Li, Liang Chen, Yunjiao Pan, and Jihui Guan. 2020. Using bottleneck analysis to examine the implementation of standard precautions in hospitals. *American journal of infection control* 48, 7 (2020), 751–756.

[24] Ronny S. Mans, Wil M. P. van der Aalst, Rob J. B. Vanwersch, and Arnold J. Moleman. 2013. Process Mining in Healthcare: Data Challenges When Answering Frequently Posed Questions. In *Process Support and Knowledge Representation in Health Care*, Richard Lenz, Silvia Miksch, Mor Peleg, Manfred Reichert, David Riaño, and Annette ten Teije (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 140–153. https://doi.org/10.1007/978-3-642-36438-9_10

[25] Francesca Marazza, Faiza Allah Bukhsh, Jeroen Geerdink, Onno Vijlbrief, Shreyasi Pathak, Maurice van Keulen, and Christin Seifert. 2020. Automatic Process Comparison for Subpopulations: Application in Cancer Care. *International Journal of Environmental Research and Public Health* 17, 16 (2020). https://doi.org/10.3390/ijerph17165707

[26] Silvia Mena-del Horno, Lirios Dueñas, Enrique Lluch, Adriaan Louw, Alejandro Luque-Suarez, Michel GCAM Mertens, Laura Fuentes-Aparicio, and Mercè Balasch-Bernat. 2022. A Central Nervous System Focused Treatment Program for People with Frozen Shoulder: A Feasibility Study. *International Journal of Environmental Research and Public Health* 19, 5 (2022). https://doi.org/10.3390/ijerph19052628

[27] Jorge Munoz-Gama, Niels Martin, Carlos Fernandez-Llatas, Owen A. Johnson, Marcos Sepúlveda, Emmanuel Helm, Victor Galvez-Yanjari, Eric Rojas, Antonio Martinez-Millana, Davide Aloini, Ilaria Angela Amantea, Robert Andrews, Michael Arias, Iris Beerepoot, Elisabetta Benevento, Andrea Burattin, Daniel Capurro, Josep Carmona, Marco Comuzzi, Benjamin Dalmas, Rene de la Fuente, Chiara Di Francescomarino, Claudio Di Ciccio, Roberto Gatta, Chiara Ghidini, Fernanda Gonzalez-Lopez, Gema Ibanez-Sanchez, Hilda B. Klasky, Angelina Prima Kurniati, Xixi Lu, Felix Mannhardt, Ronny Mans, Mar Marcos, Renata Medeiros de Carvalho, Marco Pegoraro, Simon K. Poon, Luise Pufahl, Hajo A. Reijers, Simon Remy, Stefanie Rinderle-Ma, Lucia Sacchi, Fernando Seoane, Minseok Song, Alessandro Stefanini, Emilio Sulis, Arthur H.M. ter Hofstede, Pieter J. Toussaint, Vicente Traver, Zoe Valero-Ramon, Inge van de Weerd, Wil M.P. van der Aalst, Rob Vanwersch, Mathias Weske, Moe Thandar Wynn, and Francesca Zerbato. 2022. Process mining for healthcare: Characteristics and challenges. *Journal of Biomedical Informatics* 127, 103994 (2022). https://doi.org/10.1016/j.jbi.2022.103994

[28] Andrew S. Neviaser and Jo A. Hannafin. 2010. Adhesive Capsulitis: A Review of Current Treatment. *The American Journal of Sports Medicine* 38, 11 (2010), 2346–2356. https://doi.org/10.1177/0363546509348048 PMID: 20110457.

[29] Robert J Neviaser and Thomas J. Neviaser. 1987. The Frozen Shoulder Diagnosis and Management. *Clinical orthopaedics and related research* 223 (Oct. 1987), 59–64. https://doi.org/10.1097/00003086-198710000-00008

[30] Gustavo Pereira, Eduardo Santos, and Marcell Maceno. 2020. Process mining project methodology in healthcare: a case study in a tertiary hospital. *Network Modeling Analysis in Health Informatics and Bioinformatics* 9 (Dec. 2020). https://doi.org/10.1007/s13721-020-00227-w

[31] Eman RABABAH, Hashem Abu Tariah, Raed Halalsheha, and Mohamad Abo Kebar. 2020. Frozen shoulder: Pathogenesis, Diagnosis and Treatment. *Journal of Kerman University of Medical Sciences* 27, 5 (2020), 447–455. https://doi.org/10.22062/jkmu.2020.91477

[32] Kaspar Riesen, Andreas Fischer, and Horst Bunke. 2014. Improving Approximate Graph Edit Distance Using Genetic Algorithms. 63–72. https://doi.org/10.1007/978-3-662-44415-3_7

[33] Tracy Robinson, Cate Bailey, Heather Morris, Prue Burns, Angela Melder, Charlotte Croft, Dmitrios Spyridonidis, Halyo Bismantara, Helen Skouteris, and Helena Teede. 2020. Bridging the research-practice gap in healthcare: a rapid review of research translation centres in England and Australia. *Health research policy and systems* 18, 117 (Oct. 2020). https://doi.org/10.1186/s12961-020-00621-w

[34] Anne Rozinat. 2017. How to Perform a Bottleneck Analysis With Process Mining. https://fluxicon.com/blog/2017/01/how-to-perform-a-bottleneck-analysis-with-process-mining/ [Online; accessed 02. Jun. 2022].

[35] Hubert Scheuerlein, Falk Rauchfuss, Yves Dittmar, Rüdiger Molle, Torsten Lehmann, Nicole Pienkos, and Utz Settmacher. 2012. New methods for clinical pathways-Business Process Modeling Notation (BPMN) and Tangible Business Process Modeling (t.BPM). *Langenbeck's archives of surgery / Deutsche Gesellschaft für Chirurgie* 397 (Feb. 2012), 755–761. https://doi.org/10.1007/s00423-012-0914-z

[36] Andrey Skobtsov and Anna Kalenkova. 2019. Efficient Algorithms for Finding Differences between Process Models. 60–66. https://doi.org/10.1109/ISPRAS47671.2019.00015

[37] John M. Westfall, James Mold, and Lyle Fagnan. 2007. Practice-Based Research—"Blue Highways" on the NIH Roadmap. *JAMA* 297, 4 (Jan. 2007), 403–406. https://doi.org/10.1001/jama.297.4.403