Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

INGEGNERIA BIOMEDICA, ELETTRICA E DEI SISTEMI

Ciclo 35

Settore Concorsuale: 01/A6 - RICERCA OPERATIVA

Settore Scientifico Disciplinare: MAT/09 - RICERCA OPERATIVA

ANALYTICS AND OPTIMIZATION FOR EMERGENCY HEALTHCARE PROCESSES

Presentata da: Cristiano Fabbri

Coordinatore Dottorato

Michele Monaci

Supervisore

Enrico Malaguti

Co-supervisore

Michele Lombardi

Esame finale anno 2023

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

Abstract

Department of Electrical, Electronic and Information Engineering
"Guglielmo Marconi" - DEI

PhD Programme in Biomedical, Electrical and System Engineering
(Automatic Control and Operational Research Curriculum)

Analytics and optimization for emergency healthcare processes

by Cristiano FABBRI

This thesis deals with the analysis and management of emergency healthcare processes through the use of advanced analytics and optimization approaches. Emergency processes are among the most complex within healthcare. This is due to their non-elective nature and their high variability. These characteristics make them ideal environments for the application of quantitative data-driven techniques.

This thesis is divided into two topics related to emergency processes. The first one concerns the core of emergency healthcare processes, the emergency department (ED). In the second chapter, we describe the ED that is the case study. This is a real case study with data derived from a large ED located in northern Italy. In the next two chapters, we introduce two tools for supporting ED activities. The first one is a new type of analytics model. Its aim is to overcome the traditional methods of analysing the activities provided in the ED by means of an algorithm that analyses the ED pathway (organized as event log) as a whole. The second tool is a decision-support system, which integrates a deep neural network for the prediction of patient pathways, and an online simulator to evaluate the evolution of the ED over time. Its purpose is to provide a set of solutions to prevent and solve the problem of the ED overcrowding.

The second part of the thesis focuses on the COVID-19 pandemic emergency. In the fifth chapter, we describe a tool that was used by the Bologna local health authority in the first part of the pandemic. Its purpose is to analyse the clinical pathway (taking into account the services relevant to the management of COVID-19 patients) of a patient and from this automatically assign them a state. Physicians used the state for routing the patients to the correct clinical pathways. The last chapter is dedicated to the description of a MIP model, which was used for the organisation of the COVID-19 vaccination campaign in the city of Bologna, Italy. It is powerful multi-scenario tool, which made the organisation more effective and efficient.

Acknowledgements

First, I would like to thank my supervisor Prof. Enrico Malaguti for supporting me during this journey. Despite all the difficulties and obstacles during these three years, he never let me lack his support and backing. A great person as well as a brilliant academic, he has been an inspiration to me.

I would also like to thank my co-supervisor, Prof. Michele Lombardi, one of the most brilliant scientists I have had the privilege of collaborating with. His great passion and example pushed me to discover such an exciting field of research as machine learning.

I also want to thank Prof. Michele Monaci, a key figure in my academic and professional growth over these three years. His advice and support allowed me to complete my most ambitious Ph.D. project.

Special thanks go to my mentor, Dr. Paolo Tubertini, who despite my every mistake never stopped giving me confidence and support.

I also thank my partner in misfortune Marco Leonessi for supporting me during these years of emergency. Our daily chats and collaboration have been one of the best part of the work of these three years.

Thanks also to all the other professors of the group of Operations Research of the Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi" of the University of Bologna: Prof. Valentina Cacchiani, Prof. Silvano Martello, Prof. Paolo Toth, Prof. Daniele Vigo; and thanks to all my colleagues for the time spent together: Dr. Luca Accorsi, Dr. Carlos Emilio Contreras Bolton, Francesco Cavaliere, Dr. Naga Venkata Chaitanya Gudapati, Henri Lefebvre, Dr. Federico Naldini, Alan Osorio, Antonio Punzo and Dr. Carlos Rodrigo Rey Barra.

Finally, the biggest thanks goes to my parents, who are essential in my life, and to my girlfriend Giulia who is always ready to give me her support in all my silly ideas.

Contents

List of Figures

List of Tables

List of Abbreviations

DL	**D**eep **L**earning
DNN	**D**eep **N**eural **N**etwork
DPH	**D**epartment of **P**ublic **H**ealth
DSS	**D**ecision **S**upport **S**ystem
ED	**E**mergency **D**epartment
GP	**G**eneral **P**ractitioner
HUA	**H**igh-**U**rgency **A**rea
IIS	**I**nformation **S**ystem of the **S**pecialist
LBET	**L**eft **B**efore **E**nd-of-**T**reatment
LCC	**L**ocal **C**oordination **C**enter
LHA	**L**ocal **H**ealth **A**uthority
LIS	**L**aboratory **I**nformation **S**ystem
LoS	**L**ength **o**f **S**tay
LUA	**L**ow-**U**rgency **A**rea
LWBS	**L**eave **W**ithout **B**eing seen
MIP	**M**ixed-**I**nteger **P**rogramming
ML	**M**achine **L**earning
NLP	**N**atural **L**anguage **P**rocessing
RIS	**R**adiological **I**nformation **S**ystem
SSO	**S**hort-**S**tay **O**bservation
TAH	**T**ransferred to **A**nother **H**ospital
VSC	**V**accine **S**upply **C**hain

Dedicated to Giulia

Chapter 1

Introduction

Throughout the world, the organization, management and improvement of health care systems is a major task for governments as well as a topic of particular public concern. Indeed, not only health care is a fundamental component of the welfare provided to the population, but also it is above all a fundamental right that must be guaranteed.

The organization of health services, however, is a task of extreme complexity. On the one hand, this is due to the wide variety of services that need to be made available, ranging from specialized outpatient clinics to primary care and surgery. On the other hand, complexity is related to the large amount of resources (e.g. money, people and equipment) that are needed to provide an adequate level of service. Moreover, the use of resources is likely to increase due to phenomena such as the gradual aging of the population (WHO, 2022) and the consequent increase in chronic diseases (Petrelli et al., 2021). This factor has been compounded by the COVID-19 pandemics, which has hit the world since 2019. This event not only created new challenges for health systems, both clinically (Rehman et al., 2021) and organizationally (Mohamed et al., 2022), but also showed the danger of resource scarcity in health care. During the pandemic there was often a gradual suspension of basic services (e.g. ambulatory care services or surgical care services) in order to allocate resources for the emergency. The suspension of these services had serious consequences not only in the pandemic peak period, but also after recovery (see for some examples Griewing et al., 2022, Alkharashi et al., 2022, Dionisie et al., 2022 and Araja et al., 2022).

By taking into consideration the scenario just described, it becomes clear that the ability to provide the best services to people through the rational and efficient use of resources, which might otherwise fall short, is already a key challenge for health systems around the world. The conjugation of these aspects, which may be perceived as conflicting, could provide an opportunity to avoid difficult choices in the future, even in the event of a resumption of

COVID-19 pandemic emergency or in a similar situation.

The capability of successfully tacking such a challenge also comes through the adoption of approaches, methodologies or disciplines that enable the development of evolved decision support systems. Recent years have increasingly seen the emergence of new techniques (e.g. Process Mining) or the growing adoption of existing ones (e.g. Data Mining and Machine Learning) based on data (Analytics) and supporting traditional Process Management techniques (e.g. Lean Management or Agile project management). The latter, although providing a number of qualitative general purpose solutions that are effective in the general context, tend to be less effective on the particular and high dynamic cases, such as the health care processes. In addition to analytical techniques, which allow for the quantitative and objective analysis of problems, the use of operations research techniques (e.g. combinatorial optimization or simulation) allow us to transform the obtained data into insights for better decision making.

To conclude, the correct combination of these techniques is a recipe for building decision support systems (in each case subject to the control of the human decision maker) that will enable the design and improvement of health care systems, thus creating the right balance between quality and cost of service.

1.1 Emergency and urgency services: a brief description

The health care processes and services provided within the various systems differ one from each other. These differences may relate to multiple factors such as the type of professional required, acceptable delay before the service is provided, or the type of target patient. Regardless of other characteristics, the first level of classification of health care processes concerns the difference between elective care and non-elective care (Mans et al., 2015).

The first type covers all activities, services and processes whose planning can be done in advance, involving patients who need care that can be delayed (Gupta and Denton, 2008). Ambulatory care services, surgical care services (the share of non-urgent surgery) or home care services can be found within this group.

The second type deals with all those services dedicated to patients with unexpected needs that cannot be organized in advance, but only at the time of their onset (Gupta and Denton, 2008). The most important processes, of this second group, concern the emergency and urgency services. They affect a large number of networks, organizational systems, and professionals within emergency medicine, identified as the medical specialty dedicated to treating, managing, and preventing unexpected injuries and illnesses (Schneider et al., 1998). Patients who are treated within this specialty can be grouped within two classes: urgent and emergent. The former includes patients with problems that can be treated within a short period. The second, on the other hand, concerns patients who need to be treated immediately (Mans et al., 2015). Although several facilities contribute to make up the emergency-urgency system (e.g. ambulance networks, trauma centers or emergency surgery), the center is the emergency department (ED) (Hulshof et al., 2012).

The ED, also known as emergency room, is the facility dedicated to the treatment and care of patients with unexpected or life-threatening injuries and illnesses. Often, in addition to urgent and emergent patients, the non-urgent class of patients who do not find an adequate response within the primary care system (see Beaulieu et al., 2013) are also treated. Internally, given the non-elective nature of the processes, illnesses treated and arrivals, patients who show-up are organized through a prioritization process called *triage*. The triage is a set of procedures, usually performed by a nurse, through which a priority code (numerical or by color) is assigned to patients so that the more clinically serious ones can be seen before the less serious one. After the triage, the patient is examined by an ED physician[1] and possibly discharged or referred for further investigations until a complete clinical picture is obtained.

Due to their nature, processes related to emergency and urgency are among the most complex to manage within health care systems. The inability to plan activities in advance means that there is a risk of under-utilization or over-utilization of the resources, resulting in waste or poor service provided to the patients. This issue makes this subject area the ideal candidate for the application of advanced quantitative analytical and operational methods.

[1]This phase is known as visit or first visit.

1.2 A special case of emergency: the COVID-19 pandemics

Since 2019 a different kind of emergency has spread all over the world, namely the COVID-19 pandemic. The infection was first identified in the Chinese city of Wuhan in December 2019 (Lu et al., 2020), but soon spread to almost every country in the world (Chen et al., 2020, C. Wang et al., 2020, Huang et al., 2020). Later, on January 30^{th} 2020, the COVID-19 outbreak was declared by WHO as the sixth public health emergency of international concern, following H1N1 (2009), polio (2014), Ebola in West Africa (2014), Zika (2016) and Ebola in the Democratic Republic of Congo (2019) (Lai et al., 2020), and on March 11^{th} 2020 it was declared a pandemic. To deal with the emergency and contrast its advance, in addition to using standard protocols, many governments around the world have decided to adopt increasingly harsh solutions, from suspending all nonessential medical services (especially elective ones) and the facial masks utilization, to the measure of lockdown (Onyeaka et al., 2021). Independent of the clinical and organizational solutions implemented by different countries, the most effective system in stopping the pandemic has been vaccine development. The first versions passed clinical trials at the end of October 2020 and then were administered to an increasing number of people, enabling the resumption of all previously suspended clinical, work, and social activities (Hadj Hassine, 2022).

In its early stages, the processes to face the pandemic had all the characteristics of emergency processes. Activities were purely non-elective, with uncertainty about arrivals, both in terms of time and numbers. Added to this was the uncertainty about the correct clinical and organizational model to be used to deal with this emergency, as it was different from those faced in the recent past. The main goal at this stage was to find the best solutions, at different levels, to slow the spread of the virus as it evolved, without any possibility of planning ahead.

Subsequently, with the production of the first vaccine doses, the focus shifted to organizing a mass vaccination campaign as quickly as possible. This issue, unlike the previous one, takes on the characteristics of organizing elective activities, such as ambulatory care services. The goal of this phase was to produce plannings that were as effective and efficient as possible, since increased vaccination coverage would have allowed normal activities to resume sooner.

In conclusion, although the services in order to contrast the pandemic was for all intents and purposes managed as an emergency ones for most of its duration, with the production of vaccines the problem shefted to the management of an elective process. Again, the use of advanced tools enabled better management of both phases.

1.3 Aims and objectives

The purpose of this thesis is to propose hybrid analytics and optimization tools to support the activities of professionals (physicians, nurses or health workers) working in emergency health systems. All the tools described in the following chapters have been applied to real cases, and most of them have been used as real support for practitioners.

In order to provide a clear methodological reference on which area each tool that will be described covers, two theoretical models need to be referenced. The first is the taxonomy proposed by Hulshof et al., 2012, in which the authors combine health care application areas (e.g. emergency care service, ambulatory care service) to decision-making and planning levels (e.g. tactical planning, operational planning). The second model is the Analytics Maturity Models, in which, depending on the depth of data use by an organization or tool, there are different impacts and costs (from Descriptive Analytics to Cognitive Analytics) (Król and Zdonek, 2020).

The thesis is divided into two case studies, which have been the subject of three years of work. Two tools were developed for both of them, in order to support the activities. The first concerns a case study from a real ED, located within the Italian territory. The second concerns the COVID-19 emergency with reference to the context of the metropolitan city of Bologna in Italy.

1.3.1 Emergency Department case study

Chapter 2 The case study for which the two tools covered in chapters 3 and 4 were developed is presented in Chapter 2. First, a conceptual model of ED operations is described, which is fundamental to understanding the problem. Next, a retrospective analysis of the ED and its internal organization is provided.

Chapter 3 In Chapter 3 we propose a new type of diagnostic analytics model. We start by explaining why traditional analytics systems tend to view ED pathways and their support services (e.g., radiology) as separate elements.

This fact, instead of providing the appropriate support, can lead to incorrect consideration and a bad allocation of resources. The proposed tool, which takes inspiration from simulation models, tries to overcome this critical issue by analyzing the process as a whole.

Chapter 4 In Chapter 4 we propose a new Decision Support System (DSS) based on the integration of a Deep Neural Network (predictive analytics) and an online simulation (prescriptive and cognitive analytics with OR) tool applied to the ED (Online operational planning). The purpose of this system is to provide valuable support for the abatement of overcrowding, possibly before it occurs. The neural network is used to predict patients' clinical pathways by exploiting all information collected during the triage process. Predicted pathways are used within a discrete-event simulator in order to test different policies and to select the most appropriate one so as to decrease overcrowding.

1.3.2 COVID-19 case study

Chapter 5 In Chapter 5 we present the organization of the response to the pandemic emergency within the Italian city of Bologna. We describe the patients' taking charge process and introduce an algorithm that gives them a state, through which they are directed to the appropriate path (prescriptive analytics for offline operational planning level). This tool gave a fast automatic way to manage the system and to define the best clinical strategy. The analysis showed that the use of such tool helped healthcare professionals during the most difficult moment of the emergency (the first 6 months), when swabs were scarce and information systems were still not adequate.

Chapter 6 In Chapter 6 we present the organization and scheduling of a vaccination campaign during a pandemic emergency. We describe the decision process and introduce an optimization model, which showed a powerful multiscenario tool for scheduling a campaign in detail within a dynamic and uncertain context (tactical planning). The solution of the model gave the decision maker the possibility to test different settings and have a configurable solution within few seconds, compared with the man-days of effort that would have required a manual schedule.

Chapter 2

Emergency department case study

This chapter introduces the case study which is the object of the next two chapters. The proposed approaches have been tested on a data set derived from real observations from one of the largest metropolitan ED of a northern Italian region. The considered ED is classified as a second-level emergency acceptance department (DEA 2), the highest complexity level within the Italian classification (see Italian Ministry of Health, 2008). Such an ED provides a number of highly specialized services, and a 24/7 radiology.

In this chapter, after an introduction dedicated to the description of a generic ED organizational model (Section 2.1), both the internal organization of the ED (Section 2.2) and the historically relevant data (Section 2.3) of the case study are described.

2.1 Introduction

All EDs include the same kind of human actors and resources such as doctors, nurses, clinical staff, technicians, devices, beds and stretchers, all of them interacting within similar processes. Therefore, in this section we present a general model for describing an ED path and the interaction with the resources. Although the case study under examination comes from an ED located in Italy, the two approaches that will be presented during the next two chapters are general and easily adapted to other contexts.

A patient within an ED follows a path that can be summarized as in Figure 2.1.

Each patient enters the ED either autonomously or with an ambulance. In both cases, the *triage* phase starts as soon as possible, and a staff member, usually a nurse, assesses the patient and registers their personal information (name, age, sex, etc.) and clinical observations (oxygen saturation, blood pressure, etc.). In addition to the structured information, nurses often fill

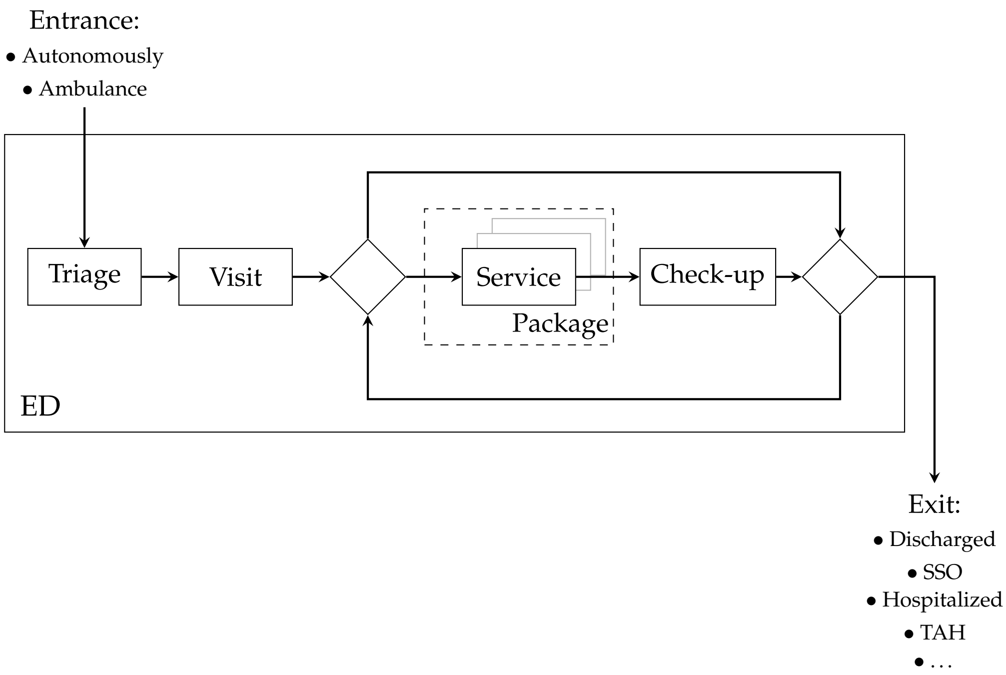


FIGURE 2.1: General ED path

in a text box (nursing diary) with more detailed information about the patient, such as the type of injury or other diseases affecting the patient. This unstructured textual information is often more relevant than the structured one. For example, in the ED we considered, the nursing diary was filled for every patient registered in the period from January 2018 to October 2019; in the same period, the field "Main Problem" recorded a very generic information ("other injury") in more than 50% of the cases. After receiving a priority during the triage phase, each patient is possibly placed in a waiting room. In case the waiting time extends too much, as it may happen in overcrowded EDs, some patients may leave without being seen (LWBS). After the waiting, a patient is checked for the first time by a physician (Visit).

At the end of the visit, the physician either decides to discharge the patient or requests a further set of services, including, for example, X-ray exams, ultrasound, specialist visit, laboratory exams, therapies etc. In the following we will denote by "package" a set of services which are prescribed together for the same patient without a predefined order. A check-up is performed after each package is completed; during this check, the patient is reevaluated and the physician either discharges them or requests an additional package. This loop can be repeated several times, until the physician has a diagnosis.

Once patients receive the diagnosis, they may have different destinations:

discharged, hospitalized, transferred to a Short-Stay Observation (SSO) area, or transferred to another hospital (TAH). In the last three cases, the bed availability has a major impact on the patients' length of stay (LoS).

2.2 Internal organization

In this section we present how the case study ED is organized.

After arrival, patients undergo the triage process, which is performed by two nurses working in parallel. They can be helped by one more nurse during period of overcrowding and who normally works as a flow manager. During this phase, low-urgency patients with particular needs (e.g. eye problems) are enrolled in pathways called fast-track, and sent to a specialist.

After triage, patients await the visit in two rooms, outside of the ED area, depending on whether or not they need stretchers. Then, depending on the urgency, patients (not LWBS) can be visited in an high-urgency area (HUA) or in a low-urgency area (LUA).

The first one is organized as an open-space area with 8 stretchers and manned by 2 physicians, 3 nurses and 1 health worker who deals with the movement of patients. This type of organization facilitates the monitoring of multiple patients who need to be treated for a long time (tens of minutes), after the visit itself. The area is always open and during the night (from 8PM to 8AM) receives also low-priority patients, due to the closure of the LUA.

The second one is composed of 4 ambulatories served by 2 physicians, 2 nurses and 1 health worker who deals with the movement of patients. This type of organization allows staff to change ambulatory at the end of a visit (or check-up), so that it can be restored. In case of a very long queue of high-urgency patients, they can be also admitted to these ambulatories. The area is open from 8AM to 8PM.

In the event of the arrival of a severely-urgent patient, one physician and one nurse are moved to a shock-room, where they can provide special care.

In addition to the other physicians, there are two other physicians who work within ED in order to support special activities. The first one is an expert physician that works between 11AM to 7PM as support to other doctors for patients with particular diseases or in case of problems. The second one works 6PM to 12AM, in order to close clinical cases after LUA closing or physicians' change shift (8PM).

When patients cannot be discharged (e.g. they need therapy), but it is not necessary that a physician monitor them or they have to wait for other services, they are moved to a waiting room within the ED. It is checked by a nurse 24 hours a day.

In the area near to the ED, there are 3 radiology clinics, dedicated to the ED, where patients who require X-Ray exams or ultrasounds are accepted. In an other area of the hospital there are 1 MRI scan and 1 CT scan, which are not dedicated, but the ED takes precedence.

Concerning laboratory exams, the samples of blood are taken from the patient and sent by pneumatic mail. Physicians and nurses receive the results through an IT system.

Moving to specialist visits, physicians are available for the ED 24 hours a day (in presence during the day and through calling during the night). The specialist checks the patient within the ED or in an other area, depending on the discipline.

At the end of the pathways or a pack of services, the physicians checks patients in most of cases within the LUA or HUA.

Once a patient has completed their urgency treatments and has received a diagnosis, they may have different destinations:

- discharged, in case the there is no immediate need for further clinical treatments or assistance;
- hospitalized, in case further complex clinical treatments are needed;
- transferred to a SSO, in case short stay observation and/or low intensity care is needed;
- TAH;
- other possible results (including death).

2.3 Retrospective data analysis

In this section we analyze the available historical data from the ED at study. We considered anonymized real information registered for patients treated between January 1^{st} 2018 and October 31^{st} 2019, for a total of 109,201 accesses [1]. Table 2.1 shows the number and percentage share of patients for each of the 5 urgency codes, from the most to the least urgent.

Table 2.2 reports the different kind of pathways of each patient who entered the ED, disaggregated by urgency code. Accordingly, we do not report

[1] all data were treated according to the European GDPR 2016/679

Urgency Code	Type of patients	#	%
1	Life-threatening patients	6193	5.7
2	Patients who need an urgent visit	27249	25.0
3	Patients who need a not-urgent visit	56381	51.6
4	Patients with minor injuries	19301	17.6
5	Patients who died before arrival	52	0.1

TABLE 2.1: Urgency codes: description, number and percentage share.

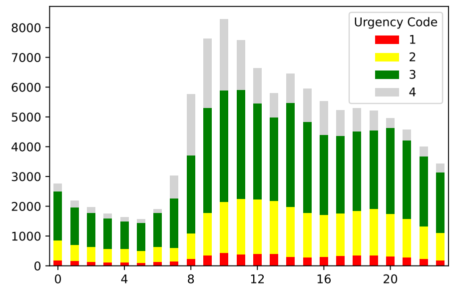
figures for patients who died before arrival. The table shows that most of the patients follow the normal pathways, with the noticeable exceptions of those following the fast-track path or leaving the ED before being seen (LWBS). These patients are not further considered in our analysis, since they do not use the ED resources and leave it after the triage. Therefore, we focus on the remaining 81,771 accesses (49%Female and 51%Male), with an average age of 55.6, which include patients who had normal pathways and who left before end-of-treatment (LBET); for the considered patients, we registered an average of 122.4 daily accesses, with a median of 123, a minimum of 86 and a maximum of 153.

Treatment Urgency Code	Normal pathways	LBET	Fast-Track	LWBS
1	6193	0	0	0
2	26879	174	0	196
3	45767	282	94	10238
4	2419	57	12011	4814
Sum	81258	513	12105	15248

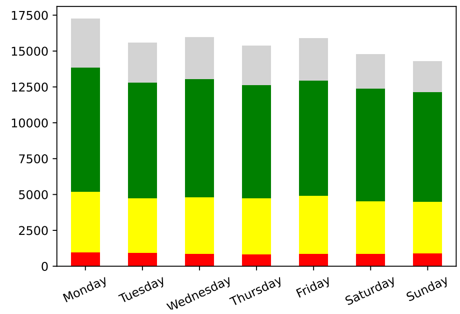
TABLE 2.2: Type of treatment based on urgency code

Figure 2.2 reports the daily (2.2a) and weekly (2.2b) trend of arrivals for each urgency code. The figure shows that, while the trend of arrivals for each urgency code is almost independent on the day of the week, it is strongly affected by the hour of the day, with a peak in late morning, and 68% of arrivals between 8AM and 8PM.

The main symptoms registered for the patients at triage are shown in Table 2.3. In the majority of the cases, the nurse selected a generic “Other injuries” entry, and provided more unstructured information in the nurse diary. This observation witnesses the importance of the nurse diary as a source of valuable information for forecasting the patients’ pathways.



(A) Daily trend of arrivals



(B) Week trend of arrivals

FIGURE 2.2: Trend of arrivals

Symptom	Frequency
Other injuries	50.5
Trauma	17.2
Abdominal pain	8.7
Dyspnea	5.1
Chest pain	4.0
Stroke	3.1
Different from the previous	11.4

TABLE 2.3: Frequency of main symptoms

Table 2.4 shows the destination of patients after the ED pathway. From the data it is possible to note that as the urgency increases, the percentage of hospitalized or SSO also increases.

Urgency Code Destination	1	2	3	4	Sum
Discharged	409	9424	32134	2199	44166
Hospedalized	2673	6092	4307	60	13132
LBET	0	174	282	57	513
Other	1100	2115	1700	67	4982
SSO	1851	8462	7142	83	17538
TAH	160	786	484	10	1440

TABLE 2.4: Type of patients' destination after ED pathway based on urgency code

Finally, Table 2.5 reports, disaggregated by urgency code, the average and median waiting time for the first visit, and length-of-stay within the ED. The waiting time for high-priority-code patients (especially urgency 1) is partially overestimated due to poor quality of the data. Often, due to serious clinical conditions, this type of patients is admitted immediately and the start of the examination is registered at a later stage.

	Waiting time [min]		LoS [min]	
	mean	median	mean	median
Urgency Code				
1	11	8	154	124
2	34	19	219	191
3	210	183	361	334
4	276	230	373	326
Total	139	77	298	263

TABLE 2.5: Patients' waiting time and LoS

Chapter 3

An algorithm for ED pathways classification

In this chapter we present a new way to assess the ED performance. We start with the description of the traditional analytics system and we show why it can lead to a bad comprehension of the problems. Therefore we propose a new algorithm, derived from the discrete event simulation paradigm, in order to more effectively identify problems (e.g. bottlenecks) within the ED. The analysis of the real case study, described in the previous chapter, shows how this new technique provides to the stakeholders a clear vision of the improvement areas.

3.1 Introduction

In recent years, EDs have been receiving more and more attention from governments and the public. EDs, in fact, are one of the main gateways to the healthcare system. Their mission is the management of emergency and urgency treatments. Due to the impossibility of predicting arrivals, the non-elective nature of the patients and the severity of the diseases treated, the organisation of activities within the EDs is among the greatest challenges of a healthcare system.

Another source of complexity comes from external activities. Often the services provided in the EDs are not sufficient to clinically assess the patient, so it is necessary to refer to external services. This is the case for services such as radiology and laboratory, which are widely used in the EDs.

The combination of the described elements contributes to an environment which is complex to assess as well as to organise. In linear processes, it is easy to understand the activities that cause slowdowns or bottlenecks (see Mans et al., 2015). However, when approaching a process in which activities overlap even partially, imputing a delay to one or the other becomes a difficult task. This criticality is exacerbated by traditional reporting and analytics systems that tend to evaluate on the one hand global EDs times without assessing the real impact of external services, and on the other hand external services as if they were stand-alone and not part of a pathway. Looking at external activities outside the pathway context leads to overestimation of their duration and leads to looking externally for the causes of long patients' LoS. A wrong understanding of the phenomenon has the effect of a sub-optimal allocation of resources for improvement. For example, having incorrectly determined that an activity frequently requested by the ED has a long lead time, stakeholders will focus on reducing it, resulting in less improvement than expected.

In order to overcome this class of problems, an algorithm will be proposed in this chapter to assess the real impact of activities on the EDs pathways, while taking into account their overlap. This algorithm will provide a new perspective on the study of blocking activities within a process. The chapter is organized as follow. In Section 3.2 we report a literature review that analyze the contribution on reporting and indicator systems. In Section 3.3 we propose an algorithm for dividing and classifying an ED pathway. In Section 3.4 we present the results obtained by applying the algorithm to the case study described in Chapter 2. Finally, in Section 3.5 we conclude the chapter by proposing some possible developments of the work.

3.2 Literature Review

Process performance indicators are a fundamental part of the ED management. Through their proper development, monitoring and evaluation, it is possible to effectively respond to issues that may arise, often even before they do.

This issue has been widely addressed in the literature. Studies can be collected into two main branches: qualitative and quantitative tools.

Regarding the first stream, a remarkable approach, although not strictly related to the ED environment, was proposed by Kaplan, Norton, et al., 2000. In this work authors proposed a strategy map for building a strategic objective of an organization, in order to enhance the balanced scorecard.

In recent years, however, there has been mainly the spread of quantitative methods. This growing adoption is due to the increasing availability of data and methods based on it (Dumas et al., 2013). In addition, the development of Event Log based techniques (see Van der Aalst, 2016), such as Process Mining (see Mans et al., 2015), has made it possible to break away from traditional step-by-step analysis and focus on the whole process.

One among the most recent and noteworthy contributions of this area is Cho et al., 2020. In this paper, the authors suggest multiple process performance indicators for an ED which can be analyzed by using event logs and which are based on the devil's quadrangle, i.e., time, cost, quality, and flexibility (see Dumas et al., 2013). These were applied to a real case study from a tertiary hospital in Korea. Another paper that addresses the issue is Rojas et al., 2019, in which Process Mining is used to determine which activities, sub-processes, interactions, and characteristics of episodes explain why some episodes have a longer duration.

The problem of resource consumption and related metrics is addressed in Müller et al., 2021, where authors execute a performance analysis and evaluate quality improvements within an ED.

Outside the literature closely related to ED environment, the problem of measuring process performance has been widely addressed from a quantitative perspective. Starting from Kueng, 2000, where authors suggest to evaluate business performance through an holistic process performance measurement system, many other experiences have been proposed in many systems. For example metrics applied to supply chain were proposed in Gunasekaran and Kobu, 2007, other concerned banking environment in Wu, 2012 or service in Vom Brocke, 2007.

As shown, the problem of metrics and monitoring systems is felt not only within the ED, but also in many other contexts. In this chapter, a quantitative system will be proposed, which will not only provide an accurate indication of the elements of the ED process, but also be easy for non-technical stakeholders to read.

3.3 Traditional analytics framework

The ED is not an easy system to analyze, due to its complexity in terms of variability of the activities and clinical figures that can be involved in a patient's pathways. In addition to physicians and nurses who work within the ED, others specialists can be asked whether the patient has particular needs (e.g. radiologist or ophthalmologist). Moreover, physicians often request laboratory tests in order to check the patient's clinical parameters.

Such a system is usually assessed internally with some common KPIs, for instance:

- Number of patients during a period of time (e.g. month, year), which can be dis-aggregated by type of urgency;
- LoS, which is the time between triage and patient's discharge;
- LWBS, which is the percentage (or absolute number) of patients who leave the ED before being seen;
- Waiting time for the first visit;
- Percentage of hospitalized patients;
- Percentage of patients sent to the short-stay observation;
- National ED OverCrowding Study (NEDOCS) (see Weiss et al., 2004 for the description)
- ...

Concerning the external services, they are usually assessed separately from the ED path to which they belong to. Common measures are:

- Waiting time, which is the time between the ED physician's request and the beginning of the service;
- Execution time, which is the time to perform the service;
- Time to produce the medical report.

Although this last set of KPI is very useful for service providers, it may not be sufficient to make an accurate assessment of the ED system, especially to evaluate the actual impact on the pathway duration. This traditional approach does not take into account that the time of different services can overlap. Let us consider Figure 3.1 and suppose this scenario:

- t_0 patient starts the visit;
- t_1 physician requests a specialist visit;
- t_2 visit ends;
- t_3 specialist visit starts;
- t_4 specialist visit ends.

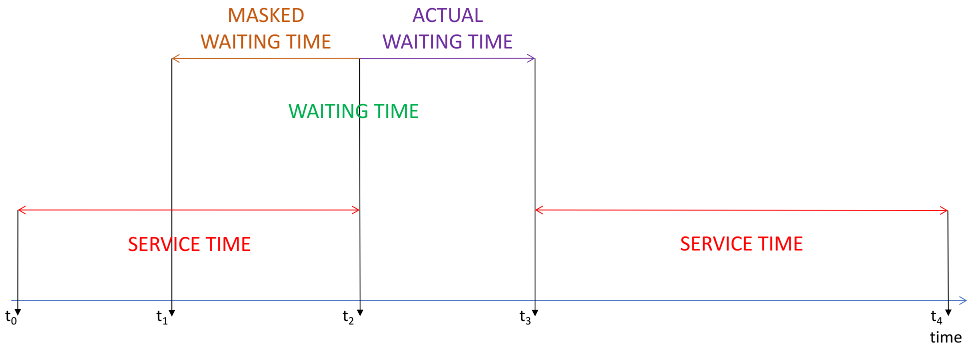


FIGURE 3.1: Example of interaction between services

With traditional analysis, this pathway is divided into two parts, with different statistics:

- ED pathway with service time $(t_2 - t_0)$;
- specialist visit, with waiting time $(t_3 - t_1)$ and service time $(t_4 - t_3)$.

The global pathway has different waiting and service time. Time from t_1 to t_2 can be considered as "*masked waiting time*", since the patient is doing something, despite they are waiting for a consultation. Therefore, if we consider a pathway as a whole the service time will be $(t_4 - t_3)$ plus $(t_2 - t_0)$, while the "*actual waiting*" time will be $(t_3 - t_2)$.

Furthermore, the services can be divided on the basis of the presence (e.g. radiology) or absence (e.g. lab tests) of the patient. In the last case it is possible to perform the activities covering the waiting time.

In the following sections we propose an alternative approach.

3.4 Algorithm description

In this section we describe the algorithm that was used for the analysis. As described in the previous section, evaluating service times, without considering the pathway is useful for those who have to provide a service, but does not provide any indication of which element is really blocking within the ED. This problem has been addressed by building a system based on processes instead of activities.

3.4.1 Data sources and log

Before proceeding with the description of the tool, it is necessary to list the data sources and the related events. During an ED pathway, many different type of physicians and IT systems can be involved. Therefore, the first part of our work concerned the analysis and organization of information related to ED pathways. Data sources can be classified into four types:

- Emergency Department Information Systems (EDIS), which contains information relating to activities carried out within the emergency department (e.g. triage information) or physician notes on the patient;
- Laboratory Information System (LIS), which contains information about the tests required for the patient;
- Radiological Information Systems (RIS), which contains information about the radiological exams required for the patient;
- Information Systems of the Specialist (IIS), which contains information about visits performed by various types of specialists.

Concerning the EDIS data source we collected these events:

Triage (*ed1*) identifies the end of the triage phase;

Beginning of visit (*ed2*) identifies the beginning of the first visit;

Discharge (*ed3*) identifies the end of the pathway.

Moving to the LIS, it is possible to identify:

Laboratory request (*ed4*) identifies the moment in which the doctor makes the request for laboratory tests[1];

Production of the laboratory report (*ed5*) identifies the moment when the doctor receives the test results.

Concerning the RIS, we can observe:

[1]This timestamp is very close to the moment of taking the sample.

Activity request (*ed6*) identifies the moment in which the doctor makes the request for a radiological activity;

Beginning of activity (*ed7*) identifies the moment in which the activity starts;

End of activity (*ed8*) identifies the moment in which the activity ends;

Production of the medical report (*ed9*) identifies the moment when the doctor receives the activity result.

Finally, the events that can be associated with the IIS are:

Activity request (*ed6*) identifies the moment in which the doctor makes the request for a specialist activity;

Beginning of activity (*ed7*) identifies the moment in which the activity starts;

End of activity (*ed8*) identifies the moment in which the activity ends.

Unlike the RIS, the instant of production of the medical report is not recorded in the IIS.

This types of events can be easily extended and adapted to specific needs (e.g. take into account the time between a lab request and sample check-in).

All of these events were organized into an Event Log, which is a very useful structure for process analysis (see Van der Aalst, 2016). The structure we have obtained is the following:

- Case_ID, which is a patient identifier within the pathway;
- Event, which is the type of event that occurred to that patient, among those described so far;
- Activity, which is the activity itself to which the event refers (e.g. Radiological);
- Timestamp, which is when the event occurred;
- Place, which is when the event occurred;
- User who performed the activity.

The last two columns are not useful for the algorithm itself

3.4.2 Limitations and assumptions

Working with data derived from real cases, some problems often arise that have to be taken into consideration, so as not to affect the quality of the analysis conducted.

One of the biggest problems is the lack of data. This can happen both due to the inability to record some events, and when, although this possibility exists, it is not exploited. Concerning the first type and the case study, we have identified four main missing data: the beginning of the triage phase, the end of the first visit, the checks-up (see Figure 2.1) and the administration of the therapy within the ED, which presents the initial instant but not the final one. As for the events not voluntarily compiled, we have found the boarding start timestamp and some request for a minor part of activities (about 10%), such as cardiology visit. The first one represents the moment in which the physician considers the patient's pathways complete, but this cannot be discharged because they are waiting for the bed to be hospitalized. In our case study we found that this event is reported by the physicians only in 20% of hospitalized patients and only in 2019 (in 2018 this event could not be reported). This is due to the fact that the physician in a stressful environment considers this activity to be of little value, despite the fact that boarding time often represents a substantial share of the LoS of hospitalized patients. As regards unregistered requests, this is due to the habit of making some of them through paper. This fact prevents the evaluation of the waiting time for these activities, but only the execution time of the same. For this class of problems it was not possible to make objective assessments.

Even when activities are recorded correctly, the problem arises of evaluating which activities have an impact on the ED pathways and which do not. This is the case of hospitalized patients for whom a service was requested before the end of the ED process and the result is ready later. In this case we have decided to adopt the following policies:

- Radiological or specialist activities will be considered if they are completed before the end of the ED path, as it happens that although the medical report arrives later, the response can reach the ED physician through other channels, such as by telephone;

- Lab services will be considered if they are completed before the end of the ED path.

3.4.3 Algorithm philosophy and functioning

To explain the operating philosophy of the algorithm, let us take the Figure 3.2 for example. It shows an example of a possible ED pathway and it highlights how requesting external services impacts the patient's LoS. The ED pathway is represented as a line, with parallel lines indicating the required services. The combination of the latter contribute to divide the ED path into phases, which we have classified. The figure can be described as follow:

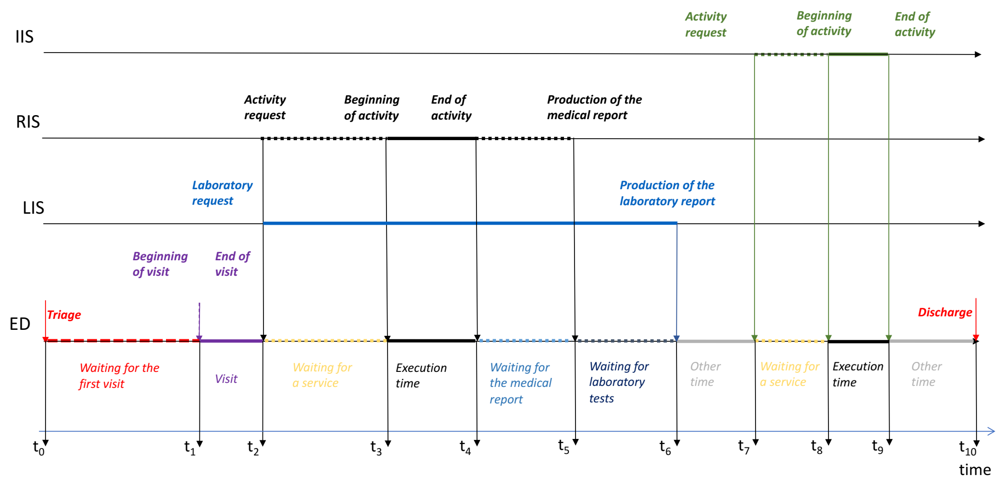


FIGURE 3.2: ED path example with the impact of external services

- The patient enters the ED and is recorded during the triage at t_0 (as previously described we do not know the exact moment of the beginning of triage);
- After a wait, the patient starts the first visit with one of the ED physicians t_1;
- At end of the visit, the physician asks for some radiological exams and lab tests t_2 (the exact moment of the end of the visit is not recorded, but after an interview with the operators it was decided to fix it at the moment of the requests for the first package of services);

- Although at this time the patient is waiting for both radiological services and lab tests, it is decided to classify this state as "*waiting for service*", since it is the radiological service that requires the patient's presence;
- For the same reason, the time elapsed between t_3 and t_4, which is the duration of the service, it is classified as "*Execution time*";
- At the end of the service, the patient has to wait for the production of the medical report until t_5 "*Waiting for the medical report*";
- Only at this point it can be said that the patient is "*Waiting for laboratory tests*, until t_6;
- We have classified the time between t_6 and t_7 as "*Other time*", that it is time that the patient spends in the ED, but the activities performed are not classified (e.g. check-ups, therapies);
- At t_7, after a not-recorded check-up, the physician asks for one more specialist service, therefore patient is in "*Waiting for service*" state until t_8 and then "*Execution time*" between t_8 and t_9;
- Finally, the patient remains in the "*Other time*" state until discharge.

The philosophy of the algorithm derives from that of discrete-event simulation, in which an entity changes its state when an event occurs.

From the above description, it is possible to see the states into which the patients' pathways can be divided. In the list below they are ordered on the basis of their priority, i.e. one state prevails over the ones below:

Visit (*st1*) in which time patient is undergoing the first visit;

Waiting for the first visit (*st2*) in which time patient is waiting for the first visit;

Execution time (*st3*) in which time patient is undergoing the treatment;

Waiting for a service (*st4*) in which time patient is waiting for at least one service service;

Waiting for the medical report (*st5*) in which time patient is waiting for at least one medical report (it is possible only for radiological services);

Waiting for laboratory tests (*st6*) in which time patient is waiting for some lab test;

Other time (*st7*) in which time patient is within the ED, but if they are doing something it is not classified.

This order can be easily modified or enriched and adapted to specific needs.

3.4.4 Classification algorithm

In this part we show the pseudo-code of the algorithm that process the logs and defines the value of the state variable. Each state change is logged in order to evaluate the system performance, even if two events occur in close moments.

Algorithm 1 Algorithm to classify ED patients' pathways

Input:

- Event log, event types and activities[2] described in subsection 3.4.1 and within *LOG*;
- Patient's last state variable as *state*;

States:

- State variables described in subsection 3.4.3;

Counting variables:

1. *NES*, which represents the number of services that the patient is doing;
2. *NWS*, which represents the number of services that the patient is waiting for;
3. *NRS*, which represents the number of medical reports that the patient is waiting for;
4. *NLR*, which represents the number of lab activity requests that the patient is waiting for.

Let us define the operation of state comparison as follows:

- $state1 == state2$ if and only if $state1$ is equal to $state2$ and vice versa;
- $state1 < state2$ if and only if $state1$ assumes a value with smaller priority (see section3.4.3) than $state2$;

[2]Activity types for our tool are divided into: radiological, specialist and others.

- $state1 > state2$ if and only if $state1$ assumes a value with greater priority (see section3.4.3) than $state2$;

The Algorithm 1 describes the classification process. Starting from the triage event, the algorithm tries to assign a specific state when the right event occurs. Basically the state changes every time an event connected to a state with higher priority occurs, or the requests pending in the current state end. The algorithm iterates until the discharge event occurs.

Algorithm 1

```
Input: state := void, NES := 0, NWS := 0, NRS := 0, NLR := 0,
cnt := 0, event := LOG.event[0], activity := LOG.activity[0]; ordered set:
LOG
while event! = ed3 do
    if (event == ed1) then
        state := st2;
    else if (event == ed2) and (state == st2) then
        state := st1;
    else if (event == ed6) then
        NWS := NWS + 1;
        if (state == st1) or (state < st4) then
            state := st4;
    else if (event == ed7) then
        NWS := NWS − 1;
        NES := NES + 1;
        if (state! = st3) then
            state := st3;
    else if (event == ed8) then
        NES := NES − 1;
        if (activity == Radiological) then
            NRS := NRS + 1;
        if (NES == 0) then
            if (NWS! = 0) then
                state := st4;
            else if (NRS! = 0) then
                state := st5;
            else if (NLR! = 0) then
                state := st6;
            else
                state := st7;
    else if (event == ed9) then
        NRS := NRS − 1;
        if (NRS == 0) and (state == st5) then
            if (NLR! = 0) then
                state := st6;
            else
                state := st7;
    else if (event == ed4) then
        NLR := NLR + 1;
        if (state == st1) or (state < st6) then
            state := st6;
    else if (event == ed5) then
        NLR := NLR − 1;
        if (NLR == 0) and (state == st6) then
            state := st7;
    cnt := cnt + 1
    event := LOG.event[cnt];
    activity := LOG.activity[cnt];
```

3.5 Results

In this section we present the results obtained by the application of the algorithm.

The test was conducted on the population described in Chapter 2, considering the pathways up to April 2019. To make the analysis more accurate and centered on the information needs of ED stakeholders, we have decided to divide the population as follows:

- With laboratory tests, i.e. patients who have at least one laboratory test within the pathway;
- With services and without laboratory tests, i.e. patients who do not have laboratory tests within the pathway, but they have at least one service;
- Without any services or lab tests, i.e. patients who do not have laboratory tests or services.

This grouping was conceived by following an interview carried out with ED stakeholders. They wanted to know how the lab time affected the ED pathways, because they thought this was the biggest problem in low-priority patient' LoS. As for the other services, nurses and physicians stated that they did not represent such a serious problem as they were requested less frequently and had less temporal impact. In support of this thesis the following aggregated data have been shown (Table 3.1). To evaluate how that time

Discipline	Statistics	Value [min]
Laboratory	Average waiting time for test results	93.2
Radiology	Average waiting time for services	28.6
	Average execution time for services	13.4
	Average production time of the medical report	21.5
Specialist	Average waiting time for services	46.3
	Average execution time for services	13.2

TABLE 3.1: Aggregate data regarding the performance of the services provided to the ED

really impacts on the path of ED, we present the results obtained by the algorithm in Table 3.1 and Figure 6.4. The first one shows the average time patients spent in each state, based on their urgency and type of pathway. The figure depicts the average percentage of time patients spent in each state during their pathways.

Urgency		Patients	Statistics [min]							
			Wait. Visit	Visit	Wait. Serv.	Exec. time	Wait. M Rep.	Wait. Lab	O.T	LoS
Code	Type	count	mean	mean	mean	mean	mean	mean	mean	mean
1	With lab tests	1336	6.9	10.6	29.9	17.7	12.9	54.5	76.9	209.4
	With services and without lab tests	2549	5.8	9.9	27.9	19.4	12.7	-	63.7	139.4
	Without any services or lab tests	192	9.1	27.5	-	-	-	-	59.2	95.7
2	With lab tests	9624	32.0	15.8	39.9	14.5	11.4	54.7	88.4	256.8
	With services and without lab tests	9164	26.1	16.3	42.6	16.6	12.7	-	83.6	197.9
	Without any services or lab tests	2255	30.4	26.3	-	-	-	-	78.7	135.4
3	With lab tests	11919	195.1	21.8	34.0	11.7	8.4	60.4	87.1	418.6
	With services and without lab tests	13747	185.6	21.1	38.5	13.5	10.7	-	84.0	353.5
	Without any services or lab tests	8019	206.1	27.7	-	-	-	-	39.3	273.1
4	With lab tests	242	278.9	27.2	39.2	12.7	5.6	66.5	91.7	521.8
	With services and without lab tests	844	171.4	21.5	27.9	11.6	5.5	-	80.8	318.7
	Without any services or lab tests	901	267.8	23.9	-	-	-	-	22.3	313.9

TABLE 3.2: Average time patients spent in each state, based on their urgency and type of pathway

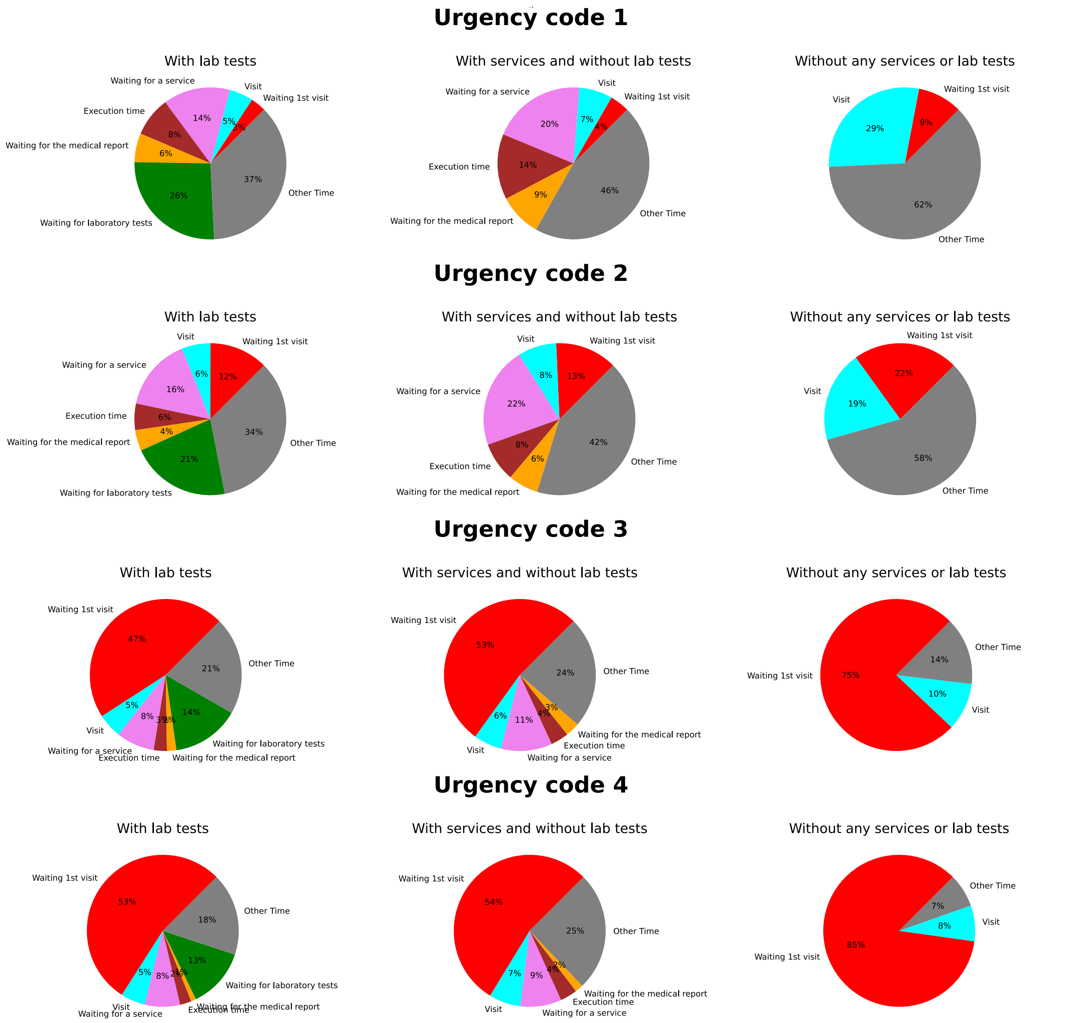


FIGURE 3.3: Average percentage of time patients spent in each state during their pathways, based on their urgency and type of pathway

As expected, as urgency decreases (from 1 to 4), both the absolute and relative impact of waiting for the first visit increases. This component becomes prevalent starting from priority 3.

Analyzing paths that contain laboratory tests, it can be seen that the impact of the laboratory is almost identical in absolute terms (from 54.5min for high-urgency patients to 66.4 for low-priority patients). These values are far from those proposed in the Table 3.1, from -42% for urgency 1 to -29%. Regardless of the urgency, the time spent in this state is not the biggest after the wait on the first visit. This share is in any case represented by the other time, which also coincides with the prevailing share for urgency 1 and 2. Other states can be considered as minority states, except for urgency 1.

Pathways with some services and without lab tests have the other time as the prevailing state (from 48% for urgency 1 to 25% for urgency 4). This is numerically similar to what is shown for the previous sample. Obviously the relative weight of other states has increased. An interesting observation is that the mean LoS between the two samples considering same-urgency pathways is the same except for the contribution of the laboratory. This is an empirical demonstration of the correct functioning of the algorithm.

As far as the last type of itinerary is concerned, there is still a prevalence of the other time, but numerically lower. At the same time the weight of the visit is increased. This fact can be partially explained by the method used to calculate the end of visit. As explained, the visit end timestamp not being recorded, is assumed as the instant of sending requests or the start of activities. Since these paths are without activity, the algorithm is led to report the end of the visit later.

From the data it appears that the share that contributes to increase the LoS is to be found within the organization of the ED (waiting for the first visit and other time) and not in external services. Despite this fact, the share connected to the laboratory can be relevant.

Concerning the performance, we implemented the algorithm in the Julia language 1.0 (see Bezanson et al., 2017), and run it on a server with an AMD Ryzen 7 2700C 8 Core processor, 64GB RAM under the Ubuntu Linux operating system. The system has been tested with a LOG composed of about 1.5M and the run lasted about 60 minutes.

3.6 Conclusions

In this chapter we have described how performance is traditionally evaluated within an ED and why this approach can lead to wrong assumptions and decisions. An ED path can be made up of many activities, provided by different professionals, which condition the LoS of the patients. The relationships between these activities are often not easy to assess, due to their possible overlap.

In order to deal with this issue, we have proposed a new algorithm which assesses the ED pathway as a whole and the actual impact of each service. The philosophy behind this algorithm is similar to the discrete event simulation paradigm. Patients' pathways were organized within an event log, so that it could be analyzed easily. We have designed a group of states in order to classify every part of the pathway and the policies by which each event triggers a state change.

The analysis performed over every ED pathways recorded between 2018 and April 2019 has showed how the biggest part of patient's LoS is spent within the ED, regardless the urgency. However, we have shown that whenever laboratory testing is needed, a significant part of the pathway depends on that activity.

Concerning future research, an evolution could focus on a better classification of the so called other time. Many events within the ED are badly recorded, so it is difficult to know how much of this patient time is due to waiting and how much to activities. Therefore, with a better work of data entry, it is possible to have even better insights about bottlenecks and other delays. Another possible evolution concerns the possibility of extending this system and the underlying algorithm to other processes not strictly related to ED. Although few healthcare processes have the same complexity as those of ED, in many cases there is a coexistence of overlapping activities. Therefore, the application of the algorithm described in this chapter could provide a better perspective on areas for improvement.

Chapter 4

On-line Strategy Selection for Reducing Overcrowding in an Emergency Department

Overcrowding is a well-known major issue affecting the behavior of an ED, as it is responsible for patients' dissatisfaction and has a negative impact on the quality of workers' performance. Dealing with overcrowding in an ED is complicated by lack of its precise definition and by exogenous and stochastic nature of requests to be served.

In this chapter, we present a Decision Support System based on the integration of a Deep Neural Network for dealing with the sources of uncertainty (e.g. patients' pathways) and a simulation tool to evaluate how specific management policies affect the ED behavior.

The DSS is designed to run on-line and to dynamically suggest the most suitable policy to be implemented in the ED. Numerical results show that a significant improvement in terms of overcrowding reduction can be obtained by allowing dynamic selection among a limited set of simple policies for queue management.

4.1 Introduction

The environment within the ED is one of the most complex in the entire health care system. This is due to both the stochastic arrivals and the non-elective nature of the activities performed. To cope with this complexity, patient admission is managed according to a priority-based policy (Leo et al., 2016), of which triage is a key part. Data collected during the triage process are useful for the whole ED pathway; as shown in Vance and Sprivulis, 2005, triage nurses are capable of assessing the patient's complexity in a reliable and valid way.

Even in case the triage correctly assigns the level of care to each patient, the performance of an ED may be affected by overcrowding (Jayaprakash et al., 2009), arising when the demand for ED services exceeds the available resources (Higginson, 2012). Overcrowding may have a negative impact on different operational aspects, such as waiting times, LoS, increasing number of patient leaving without receiving care, increasing medical errors and decreasing efficiency (Asplin et al., 2003). Overcrowding is a complex phenomenon for which there exists no universally accepted definition and measure. The most common way to quantify ED overcrowding is the so-called National ED OverCrowding Study (NEDOCS) indicator, proposed in Weiss et al., 2004. This is a one-dimensional indicator that, based on the available resources (e.g. ED beds, Hospital beds) and on the ED state (e.g. total patients simultaneously present in the ED), returns the ED overcrowding score (from "not busy" to "Dangerously Overcrowded"). Although NEDOCS is not suitable in some cases (H. Wang et al., 2014), it may be a useful indicator for detecting areas where efforts have to be put for addressing congestion (e.g., total admits in the ED, total patients simultaneously present in the ED, number of respirators, longest admit time, etc.). More in general, different Operations Research and Operations Management approaches have been proposed to face with overcrowding. A first class of actions affects the ED intake process (Welch & Savitz, 2012), including techniques for allocating ambulances within a network of EDs (Aringhieri et al., 2017), or for rerouting ambulances to others hospitals in periods of crowding (Ambulance Diversion, see Lagoe and Jastremski, 1990). A second possibility is to focus on the internal ED patient flow, in order to use available resources efficiently. Although many attempts for creating decision support system tools have been proposed in recent years (Aboueljinane et al., 2013), their application within an ED environment is challenging. Indeed, forecasting patient pathways can be hard (Rebuge & Ferreira, 2012) for different reasons, as large number of pathways variants or missing information. Nevertheless, mining the patients' pathway is a key issue for improving the internal flow of patients in the ED, avoiding waiting times or bottlenecks. The mining problem is typically addressed either using Process Mining or by means of Machine Learning. These two approaches exploit information from structured data, usually avoiding not-structured ones. Process Mining exploits data in order to provide a pathway representation (Mans et al., 2015). However, this technique tends to

be ineffective (creation of very complex models) with high variety processes (so-called Spaghetti processes), and this is the case of ED pathways. Nevertheless, in literature there are different attempts to avoid this problem. For example, the authors in Duma and Aringhieri, 2020 propose an innovative ad hoc process mining approach to discover patients' paths, that tries to solve the problem through an initial clustering of patients.
Conversely, Machine learning techniques, and artificial neural networks in particular, have the capability of predicting the future pathway of a patient (e.g Simeunović et al., 2017). This approach, unlike the previous one, makes it possible to rely on extensive information about a patient (represented via a set of *attributes*) to achieve higher accuracy predictions.

Once patients' pathways are predicted, the next step is to use this information to improve ED performance and avoid overcrowding. A commonly used approach for addressing this task makes use of simulation (Günal and Pidd, 2009; Paul et al., 2010), allowing the creation of what-if scenarios and the selection of the best resource allocation policies to improve the patients' flow. Traditional approaches use discrete-event simulation in an off-line configuration (Jun et al., 1999), to assess how a specific policy performs. Recently, alternative approaches based on agent based simulation have been introduced (Z. Liu et al., 2017).

A first attempt of taking operational decisions in an ED based on real-time prediction is proposed in Duma and Aringhieri, 2023, where an ED simulator is used to evaluate the performance of a (fixed) pre-selected policy. The implementation of this policy requires a real-time prediction of patients' pathways, which is obtained by means of a process-mining discovery model exploiting structured data. The (fixed) policy to be implemented in the real ED is then determined by evaluating a portfolio of possible policies according to some performance indicators.

Approaches based on off-line simulation and decision making are in general not fully satisfactory within a highly dynamical system such as an ED, where an on-line approach trying to solve problems before they happen could be preferable. The aim of this chapter is to propose a new DSS based on the integration of a Deep Neural Network and a simulation tool to take decision on-line. The neural network is used to predict patients' clinical pathways by exploiting all information, i.e., both structured and not-structured data, collected during the triage process. Predicted pathways are used within a discrete-event simulator aimed at *on-line* testing different policies and *dynamically* selecting the most appropriate one, so as to decrease overcrowding. In other words, the tool is designed to react immediately to any undesired behavior of the system by switching management policy when needed. We will focus on "normal" operating conditions, though the approach can be re-trained and re-calibrated to handle exceptional circumstances (such as an ongoing pandemic).

The result has been implemented and has been tested on a real case study derived from the ED described in Chapter 2. The chapter is organized as follow: Section 4.2 describes the context and data of the problem, with particular attention to the uncertainties that are addressed within the ED. Section 4.3 presents the prediction models developed to deal with uncertainties, while Section 4.4 shows the integration of simulation and optimization. Finally, Section 4.5 reports the results obtained with the DSS in our case study, and Section 4.6 summarizes and concludes the chapter.

4.2 Problem context and Data

In Chapter 2 we have described how different EDs include common features and resources. Based on the model proposed in Section 2.1, in this section we describe the uncertainties within an ED that must be taken into account within the DSS. Although the case study under examination comes from an ED located in Italy, the model itself is general and the proposed DSS can be easily adapted and used within many other EDs.

4.2.1 Uncertainty Model

In order to timely detect and react to overcrowding, one has to know how the ED will evolve in the next future. We describe the evolution of the ED system by observing its *state*, characterized by a set of measurable variables at a given time, and by forecasting the future value of those variables. As the evolution of those figures depends both on internal actions and on exogenous stochastic events, one has to deal with different sources of uncertainty, namely:

1. uncertainty on arrivals;
2. uncertainty on pathways (temporal sequence of first visit, service packages and checks-up until the patients' discharge);
3. uncertainty on duration of visits, checks-up, and services that make up the packages;

4. uncertainty on the effect of internal actions performed on the system.

Indeed, the evolution of ED state is strongly influenced by the temporal distribution of the future patients' arrivals, as well as by the patients' pathways and by the duration of each specific activity of the pathway.

Formally, let us assume we are interested in modeling uncertainty over a set of n patients, for an Emergency Department operating with m[1] possible packages (see Figure 2.1). We will proceed by introducing, for each patient i, multiple random variables and in particular:

- a random variable T_i with support $\mathbb{R}^+$, representing the arrival time for the patient;
- a random variable X_i representing the information collected on the patient at triage time. This variable is a vector of values associated with a fixed set of attributes, thus its support depends on the type of information that are collected;
- a sequence of random variables $\{Y_{ij}\}_j$ with support $M = \{1, \dots m\}$, representing the sequence of packages (i.e. the pathway) for the patient;
- a random variable D_{ijk} with support $\mathbb{R}^+$, representing the time for the k-th service, in the j-th package, for the i-th patient.

Part of our analysis (see Section 4.3) will be devoted to determine reasonable distributions and correlations for these variables.

We note that all sources of uncertainty are exogenous, i.e., they are not affected by sequencing decisions. On the other hand, the overall behavior of the ED (including performance indicators such as the number of patient waiting for a visit) depends on the complex interplay between the uncertain factors and the operated choices. For this reason, improving the performance of an ED requires to forecast these sources of uncertainty, to assess their impact, and to define how to search for an optimal policy.

4.2.2 DSS Architecture

In the following sections we will propose a DSS composed of subsystems to deal with the previous sources of uncertainty. In particular we will present:

- an arrival time generator, based statistical approaches, for the first source of uncertainty;

[1] This number takes into account one more package which is the end of the pathway

- a Deep Neural Network to predict present patients pathway;
- a service duration generator, based on statistical approaches, for the visit duration;
- a discrete-event simulator to evaluate how specific management policies affect the ED behavior.

All data-driven approaches (i.e., the statistical models and the neural network) are meant to be calibrated over available historical data.

4.3 Predictive model

In this section we present the techniques adopted to address the first three sources of uncertainty. The analysis can be applied to any ED, provided the necessary information is available.

There is major distinction in terms of predictive problems between patients that have already entered the ED (for which triage information has already been collected) and patient that might arrive in the future (for which no information has been observed). We will discuss the two cases separately.

4.3.1 Predicting pathways for patients within the ED

For patients that have already entered the ED, the most relevant aspect to be predicted is expected pathway (or what is left of it), which affects, e.g., activity queues, patients' LoS and ED global overcrowding. Thus, predicting pathways is a key issue for modelling the behaviour of the ED, though being a challenging task. Indeed, since an ED typically offers many services, the number of potential pathways for a patient is very large as the number of possible combinations of service packages grows exponentially with the pathway length. Additionally, many pathways are similar, since they include many common services (possibly, in a different order), thus increasing uncertainty in the prediction task.

From a formal perspective, this implies that the probability distribution for the package sequences (i.e. $\{Y_{ij}\}_j$) has a very large support, complex correlations, and some of the conditioning attributes (i.e. X_i) are systematically missing for specific samples. This makes probability estimation and data generation particularly complex.

In order to deal with this problem, an AI-based approach is used to forecast the actual pathways. The approach makes use of available patient's information, and predicts the packages of services for a patient one at a time (see Figure 2.1), until the whole pathway is determined. This is coherent with the metric that we adopt for evaluating the accuracy of our prediction, which considers the accuracy in predicting the *next* package, right after the current one. In addition, the proposed solution is easier that an alternative approach in which the entire pathway is forecast in one step, as the number of service packages is much smaller than the number of their combinations into pathways.

Formally, we adopt a factored approximation for the distribution of all possible package sequence. In particular, we approximate the distribution of possible sequences with a product of probabilities:

$$P(Y_{i,1}) \prod_{j=2..} P(Y_{i,j+1} \mid \{Y_{i,j'}\}_{j'=1..}) \simeq P(\{Y_{ij}\}_j) \tag{4.1}$$

In practice, when sampling the next package we use as an input to the estimator the sequence of all packages observed or generated so far for the considered patient. A specific package value (say $Y_{ij} = \perp$) is used to the denote the end of sequence.

In addition to the path taken so far (sequence of service packages), our prediction is based on a number of additional inputs/conditioning factors, i.e. X_i from our probabilistic model. These variables correspond to information that is systematically collected at triage time, namely:

- Age;
- Sex;
- Patient's urgency (triage code);
- Text in the nurse's diary.

In order to use these features within a Deep Neural Network (DNN), a preprocessing step is needed. All categorial attributes (i.e. sex and urgency) are represented using a one hot or label encoding; in our implementation, we use methods available in the scikit-learn library (Pedregosa et al., 2011). Finally, the nurse's diary consists of free text and requires preprocessing for being used. To this aim, we consider a Natural Language Processing (NLP) approach based on a Bag-of-words model. We have tested three alternatives for processing the text:

1. a combination of the Python Natural Language Toolkit (NLTK)(Bird et al., 2009) for normalizing the text (e.g., removing stop words) and scikit-learn for creating tokens and word n-grams;

2. the same approach as above, plus a stemming step, performed via the the Snowball algorithm for the Italian language (see Bird et al., 2009);

3. an open source tokenizer based on a version of the Bidirectional Encoder Representations from Transformers (BERT)(Devlin et al., 2018) for the Italian language[2] , implemented using the PyTorch framework (see Paszke et al., 2019).

A comparison of these approaches will be discussed in the Benchmark section.

All features are fed as input to our Machine Learning model, which is a Feed-forward Deep Network classifier, implemented in PyTorch. In particular we adopt a simple architecture using ReLU neurons for all hidden layers and a SoftMax activation for the output layer.

The model is trained to approximate the probability of the next package, conditioned on the information encoded by the model input. This setup it easier to sample the next package at random, with a distribution defined by the model output. A deterministic behavior can be obtained by considering the class (i.e. the package) with the highest estimated probability as the only prediction. We will consider both these operating modes in our experimental evaluation.

Known Modeling Limitations

A known limitation of our factorized approach consists in its inability to account for future packages when making predictions: correlations between such packages may arise due to hidden variable, e.g. the actual patient ailment at arrival time. However, in practice packages (i.e. groups of services) are assigned based on information that becomes available only as a result of the services themselves. For example, a medical exam may reveal additional information, which is then used to define the next package for the patient. For this reason, we expect our factorization to be a reasonably good approximation. Preliminary experiments we performed to train a model for the non-factored distribution seem to confirm this conjecture.

[2] The repository is located at
https://huggingface.co/dbmdz/bert-base-italian-xxl-cased

Moreover, in our historical dataset, prescribed packages are affected by service availability in addition to the patient condition (e.g. a particular exam or medical specialist may not be available at a specific time). Since we are not providing availability information as input to our model, these effects will act as noise on the training distribution thus making the learning problem more complex.

Lastly, it is necessary to mention how the use of machine learning techniques could be complex for an external audience (e.g. physicians). Neural networks can appear like black boxes, the results of which are difficult to understand. Therefore, their use can be perceived as an act of faith. For this reason, it becomes essential to prove their usefulness and value beyond reasonable doubt.

4.3.2 Predicting arrivals and pathways for future patients

Since no observed information is available about patient that might enter the ED in the future, we need to predict both their arrival times an their pathways.

In an emergency department, arrivals are not scheduled in advance (not elective), hence, they are best modeled as a stochastic phenomenon. Nevertheless, the large amount of historical data that is typically available allows us to forecast with good approximation the number of patients who will arrive during a specific time period.

Under the reasonable hypothesis that, for a limited time interval (e.g., one hour), arrivals are independent and occur with constant mean rate, the arrival count is well described by a Poisson process (see Bell and Wagner, 2019). An analysis of our dataset confirmed the validity of this assumption, and showed that using time-dependent rates consistently leads to better estimates. In particular, the rate seems to be driven by three factors: the hour of day, the day of week, and the month, the first having by far the largest impact. Figure 2.2a shows the average number of arrival λ_h as a function of the hour of the day h.

Accordingly, in our DSS we modeled arrivals as independent events, following an exponential distribution with time-depended rate, i.e.

$$P(T_{i+1} - T_i) = \lambda_h e^{-\lambda_h (T_{i+1} - T_i)} \tag{4.2}$$

Where λ_h is the hour of the day. By assuming $T_1 = 0$, this is sufficient to characterize the full arrival time sequence.

As a consequence of our modeling choice, the number of arrivals per hour is a Poisson process, with rate dependent on the hour of the day. This is relevant, since it allowed us to calibrate λ_h values by simply computing means over historical data. Formally, let $\{\bar{t}_i\}_{i=1}^n$ be the sequence of historical arrival times for a set of n patients; then we have:

$$\lambda_h = \frac{1}{|H|} |\{T_i \mid \text{hour}(\bar{t}_i) \in [h, h+1)\}| \tag{4.3}$$

where $\text{hour}(T)$ is the hour of the day for the arrival of the i-th patient, the $|\cdot|$ refers to the cardinality of a set, and H is the support for h, i.e. the 24 possible values for the hour of the day.

Moving to the problem of predicting the pathways, we generate those by adopting a simple statistical approach.

- First, we generate urgency code by sampling from a discrete probability distribution, that has been calibrated by computing historical frequencies for all urgency code values.
- Then, we sample packages (i.e. all $\{Y_{ij}\}_j$) recursively according to Equation (4.1), using conditional probabilities that have been calibrated via historical frequencies, computed separately for each urgency code.

The alternative would be to build a data generator for the X_i distribution, then use then neural approach. However, this approach introduces an additional source of noise. In principle, we could have introduced additional conditioning factors that are easy to determine even within a simulation, such as the hour of the day, the day of the week and the month. However, in our case study, we did not observe strong influence of these drivers on the prediction.

4.3.3 Predicting activities duration

The status of the ED also depends on the duration of services, which influences the availability of resources, the patients flow and waiting time.

We adopted a simple statistical approach to model and sample service duration. In particular, we use parameterized distributions (e.g. Normal, LogNormal, etc.), without any conditioning factor. We estimate the distribution type (from a pre-defined set) and the distribution parameters via classical approaches (e.g. sample mean and sample variance computation) over historical data. In our case we evaluated that the LogNormal distribution was the best for this purpose.

Limitations

Given the highly dynamic environment of an ED, it may happen that duration information for service is not properly registered, thus reducing data quality. Typical issues arise, for example, for services that are not registered, or are only partially registered (e.g., the staring time but not the ending time are registered), or are registered in a wrong way. These issues can be handled by means of a preprocessing step, aimed at identifying outliers, activities with incomplete information, etc.

4.4 Simulation-Optimization DSS

In this section we present the logic of our DSS: it includes a simulation tool which integrates the aforementioned predictors and can be used to test a portfolio of alternative policies for managing the ED. By selecting the best policy, the DSS performs an optimization of the ED.

4.4.1 DSS functional architecture

Each predictor of the DSS addresses a specific source of uncertainty. Those predictors are embedded within a simulator, so as to model the dynamic behavior of the ED, including patients flow, resources utilization, and the evolution of queues.

We have developed a discrete-event simulator of the ED. The simulator takes in input the actual state of the ED (hour of the day, availability and status of the resources, length of the associated queues, patients within the ED and their features) and uses the predictors to forecast the system evolution under a specific policy. The simulator functional scheme is shown in Figure 4.1. Through the DNN new packages are assigned so as to complete the pathways of the patients within the ED. Meanwhile, new arrivals are generated by the arrival predictor, and the corresponding pathways are defined through the process described in section 4.3.2. Each service of the paths is assigned a duration through the procedure 4.3.3. Both types of pathways are passed to the discrete-event simulator and used to evaluate the evolution of the ED system. The simulator was implemented within SimPy (“SimPy: Discrete event simulation for Python”, 2002), a process-based discrete-event simulation framework based on Python.

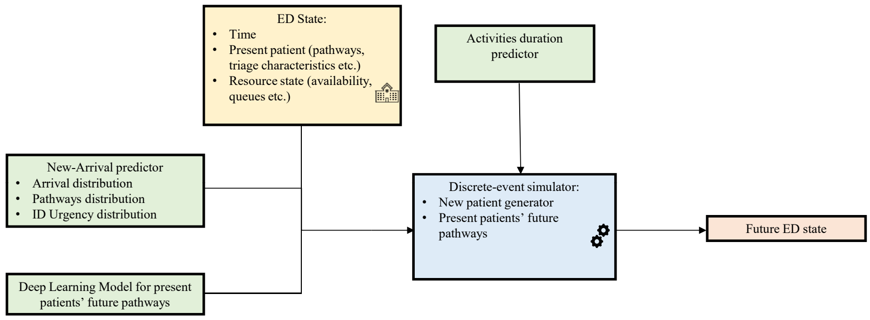


FIGURE 4.1: DSS functional scheme

The DSS is designed with the goal of identifying the most suitable policy to be implemented in the real system. To this aim, it must replicate the ED behaviour, while evaluating the impact of alternative policies. Indeed, the DSS allows the decision maker to test different policies and to choose the most appropriate one based on suitable metrics. Figure 4.2 shows integration of the DSS with the ED system. The DSS tool is triggered every Δz time units, when it is fed with the actual status of the Real-ED. It simulates the behavior of the ED under different policies, for a time interval ΔT, denoted as search depth. Each simulation is repeated ω times, so as to obtain statistically relevant information. The simulation returns, after a limited time, the selected policy to the policy-maker (red arrow), who can either accept or reject the proposal.

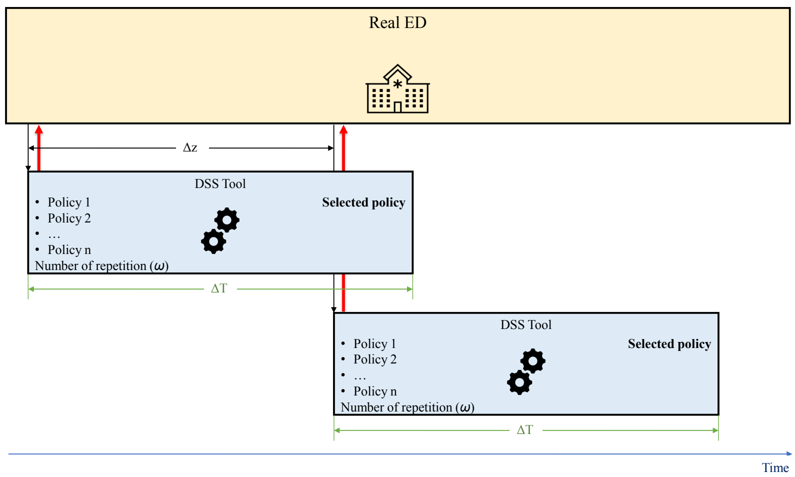


FIGURE 4.2: Integration between ED and DSS Tool

Parameters ΔT and ω affect the computing time needed to run the simulation. Clearly, having a fast answer is mandatory in order to be able to implement the suggested policy before the real ED has changed too much. In addition, using a too large ΔT value is counterproductive, as the resulting prediction is mostly based on hypotheses about future arrivals. Concerning the trigger time Δz, the lower the value of this parameter, the better the uncertainties of prediction are addressed. However, if the trigger time is too short, the implemented policy can change too often, thus confusing the decision maker.

4.4.2 Alternative policies

Since overcrowding is the major issue we want to approach by means of our DSS, when selecting the best policy to implement within the ED we consider a KPI which is a proxy of overcrowding, as detailed in the next section. Our framework is easily adaptable in case the decision maker is interested in optimizing a different KPI.

We consider alternative policies which are related to the way in which queues are handled, and select the one which performs at best with respect to our KPI. In detail, each patient can be prescribed one or more services within the current package. Services typically require the access to some scarce resource, and are associated with queues where patients wait for the resource to be available. Each patient requiring multiple services (within the same package) is assigned to the queue of the service having the shortest (expected) waiting time. The expected waiting time depends on the characteristics of other patients in the same queue, their forecast service times, and on the queue handling policy. In our DSS, we implemented three policies for selecting the next patient to be served. In all cases, priority is given according to the urgency code. In order to avoid patients starvation, the ED implements a mechanism to push forward patients who are experiencing a too long waiting time. This is obtained by increasing the urgency of a patient who has spent a certain amount of time in a queue. Ties are broken as follows:

Policy 1 selects the patient with smallest expected time for the not yet performed services of the current package, breaking ties by smallest expected time for future packages. This policy gives priority to patients who are likely to complete their current service package shortly, and can be useful in case of overcrowding.

Policy 2 selects the patient with largest expected time for the not yet performed services of the current package, breaking ties by largest expected time for future packages. This policy gives priority to patients with long expected service times and can be useful to process resource-demanding patients when the ED is not under pressure.

Policy 3 selects the next patient according to a FIFO rule. This policy is commonly used in EDs, including the one in our case study.

Regardless of the selected policy, a patient cannot change the queue to which they are assigned and moved to a different queue (before being served). This constraint follows from the design of hospitals, where ambulatories for different services can be far from each other, and moving a patient can take time.

In our analysis of the real case, we will also consider an additional policy that intentionally violates this constraint, i.e. where we assume that a patient can change their queue (in no time). Although this additional possibility is in general infeasible, its evaluation provides an optimistic estimate of the ED performance. It should be kept in mind that an implementation of such a policy could however require a redefinition of the hospital layout and logistics. We will refer to this policy as fixed-queue relaxation (FQR) policy in rest of this chapter.

4.5 Empirical evaluation

In this section we present the quantitative results of the introduced DSS and provide a detailed analysis of its capabilities, based on a real-world use case.

4.5.1 Digital twin of the real ED

The integration of the developed DSS with the real ED requires a preliminary phase of fine tuning, and a non-negligible investment in terms of time, human and financial resources. In order to perform such a tuning and assess the capability of the system before the integration is performed, we show the effect of the tool on the system through *a second simulation model*, which in our experiments plays the role of the the real system in Figure 4.2. This digital twin (DTS) is implemented as a discrete-event simulator, but one intended to be a accurate approximation of the real ED: for this reason, while it is fed with historical observations of the past ED behavior, i.e., real arrivals, patients' features (urgency, triage information etc.), clinical pathways, duration of the activities, and resource availability information.

The digital twin can be configured either with the same policy used within the real ED, so as to assess the accuracy in replicating the latter, or with different policies, so as to evaluate their impact.

The experiments performed on our case study showed that the digital twin, in the former configuration, provide an accurate approximation of the real ED. The mean LoS recorded within this system and the real one, for the whole available period (from January 2018 to October 2019), differ of 2%, and they have a similar resource occupation (global mean error 5%).

Once we have verified the good accuracy on the digital twin, we can switch to the second configuration, which is used to replace the real ED in or tuning and experiments, so as to estimate the impact of the proposed DSS.

4.5.2 Functional assumptions

In this section we describe some assumptions introduced in our analysis. These in order to manage the lack of some data and the unpredictable impact of post-ED services.

While a detailed information on the duration is available for all services in each package, this is not the case for check-ups, i.e., visits taking place right after each package of services. For this reason, based on empirical observation, the duration of each check-up is set to 15 minutes.

Although we only consider the patients' pathways within the ED, the availability of resources *beyond* the ED may have an impact on the LoS *within* the ED itself. A relevant case is bed-blocking, which can force a patient to remain within the ED after their pathway is terminated since no immediate hospitalization is possible. Our DSS has no effect on activities beyond the ED. However, in order to have a fair comparison with the real behaviour of the system, in the following analysis all figures are obtained by assuming that, after the last registered activity, each patient may have an extra waiting time in the ED depending on their pathway and final destination.

4.5.3 Performance of the Pathway Predictor

We now evaluate the performance, in terms of accuracy, of the two pathway predictors embedded within the DSS. For training and testing purposes, we considered all historical data but those from 01/10/2019 to 15/10/2019, which were reserved for evaluating the DSS.

Performance of the Predictor for Patients within the ED

Concerning patients within the ED, our historical data includes 3,704 distinct pathways. Among them, the 130 most frequent pathways cover 90% of the occurrences; these pathways are composed by 46 different packages of services. This step allowed us to considerably reduce the number of pathway variants, still covering the vast majority of the historical observations. Accordingly, the Deep Learning Classifier predicts the next package of services within these 46 variants.

In order to train, validate and test the classifier, the dataset was randomly diveded into three parts with a ratio 80-10-10. As previously mentioned, the dataset includes both structured and textual information, as well as the previously prescribed packages of services. Concerning the text processing, we used n-grams with length up to two, so as to represent semantic concepts such as "*not-something*". At the end of processing, we obtained three different vocabularies (one for each preprocessing approach described in section 4.3.1) with a different number of terms. For each considered vocabulary, we tested different configurations of the DNN in terms of number of layers, their size, and training parameters. In our study, we did not observe significant improvements by using more than 4 layers. Table 4.1 reports, for each vocabulary, the best network setup and the associated accuracy result over the test set, computed by using the predictor in a deterministic fashion (i.e. by treating as the output class the one with the largest estimated probability).

Approach Parameter	1	2	3
Vocabulary size	42034	34932	8919
Layer size	4004, 2000, 500, 46	4004, 6500, 2500, 46	6004, 3500, 1500, 46
Opt. alg.	Adam	Adam	Adam
Batch size	32	32	32
Epochs	10	3	1
Learning rate	0.005	0.005	0.005
Test set accuracy	55%	54%	62%

TABLE 4.1: Predictor results based on the NLP approach

The results show that the accuracy levels of all approaches are very satisfactory, and that Approach 3 provides the best performances. In the next section, we will show that the obtained accuracy prediction performance are appropriate and allow us to obtain a reliable DSS built on top of the predictor.

Performance of the Predictor for Future Patients

As previously described, regarding this type of patients we know nothing, then we have to make assumptions. The most important one concerns the number of access that will arrive in the next future.

As shown in Figure 2.2a, there is a very clear pattern and the hour of the day has strong predictive power. Analyzing the standard deviation in Figure 4.3, this fact is evident. Looking at figure 2.2b seems that there is a decreasing trend during weekdays, but it is rather weak.

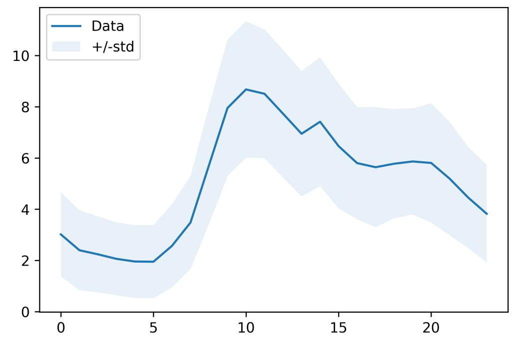


FIGURE 4.3: Arrivals mean and std. during the day

For instance we show in Figure 4.4 which Poisson distribution has a good fitting with arrivals at 2AM, 11AM, 2PM and 8PM, and after our analysis it can be extended to other hours. In order to quantify how good the distribu-

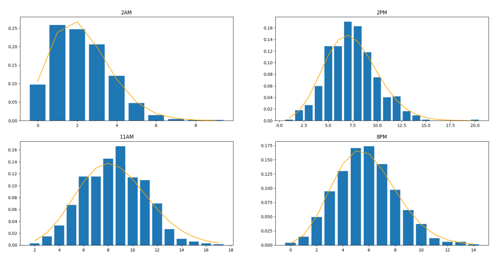


FIGURE 4.4: Poisson distribution fitting with arrival at different hour of the day

tions is, we splitted the data-set into two parts with a ratio 80-20 (training-test). We found a global (over all hours) Mean Absolute Error (MAE) over the training set of 1.64 and test set 1.68.

4.5.4 DSS performance

In this section we study the performance of the DSS tool under three metrics that are relevant for a decision-maker:

- accuracy in predicting the ED future state (number of present patients)
- accuracy in selecting the best policy over a set of possible policies
- capability of improving the ED performance

In the following, all experiments were carried out with a number of simulation repetitions $\omega = 50$. A preliminary tuning process showed that this value provides a good accuracy while requiring a short execution time (less than 10 seconds).

ED future state prevision

In order to evaluate the accuracy in predicting the ED future state, we randomly selected 100 different ED days and took snapshots of the ED state at 2AM (the time when ED has a minimum value of overcrowding), 11AM (one hour after the arrival peak, and decisions have a critical effect on overcrowding), 2PM (typically, the hour of maximum overcrowding), and 8PM (a second critical time in the day, due to the staff change of shift). These snapshots were used to initialize the ED state, and to predict the number of patients in the ED after 1, 2, and 3 hours by following the third policy described in Section 4.4.2, i.e., the one currently used in our ED. The obtained predictions were compared with the real historical values.

In Table 4.2 we report the accuracy in prediction, measured in Mean Absolute Error (MAE), between the number of predicted patients and the observed figure. Data are reported separately for each of the three approaches introduced in Section 4.5.3, both in their deterministic and stochastic version (the best configuration has been underlined in yellow).

The results show that our approaches are capable of producing tight predictions, as MAE values are smaller then their std counterparts. The best accuracy is obtained with Approach 1 in the deterministic mode, and with Approaches 2 and 3 in stochastic mode. In all cases, the performance worsens by increasing the search depth ΔT, thus showing that 3 hours is probably is a too long time period for being simulated.

Hour	ΔT	# patients mean	std	MAE Appr. 1 Det.	Stoc.	MAE Appr. 2 Det.	Stoc.	MAE Appr. 3 Det.	Stoc.
2AM	1	18.3	5.5	1.6	1.7	1.9	1.8	1.8	1.8
	2	16.6	5.7	2.7	2.9	3.0	2.8	3.0	2.9
	3	15.2	5.9	2.8	3.0	3.3	3.0	3.2	3.0
11AM	1	30.2	8.2	3.2	3.4	3.9	3.5	3.7	3.0
	2	31.5	7.7	4.5	4.6	5.9	5.0	5.7	5.8
	3	30.3	7.8	5.4	5.5	7.2	6.8	7.1	7.1
2PM	1	32.2	7.1	2.4	2.4	3.2	2.5	3.2	2.6
	2	32.6	7.4	3.9	4.0	5.3	3.9	5.2	4.0
	3	31.3	6.4	4.0	4.3	5.7	4.2	5.6	4.3
8PM	1	26.1	6.1	2.7	2.6	3.2	2.7	3.1	2.7
	2	26.3	6.6	3.6	3.6	3.9	3.9	3.7	3.8
	3	25.0	6.2	4.0	4.4	4.4	4.4	4.3	4.4

TABLE 4.2: ED future state prediction results: historical data and Mean Absolute Error of the considered prediction approaches.

Best policy prediction for improving the ED performance

The ED performance can be evaluated under different metrics whose relevance varies for different stakeholders. A first performance indicator we consider is overcrowding, for which a good proxy is given by the average number of patients within the ED. Overcrowding affects the stress level of the operators and hence the quality of the service provided within the ED. From the patient's perspective, a very relevant indicator is instead represented by the overall length of stay. Clearly, under the reasonable assumption that arrivals are independent on the ED state, these indicators are two faces of the same phenomenon: the smaller the average length of stay of patients, the smaller the average number of patients in the ED. However, while number of patients can be measured for each instant, thus being a meaningful figure even for small time intervals, the average length of stay is typically some hours; thus, its value is meaningful only when a sufficiently large time interval is considered. As the simulated interval (search depth) ΔT is much smaller than the average length of stay, within the DSS we use the average number of patients in the simulated interval, computed as

$$c_{avg} = \frac{\int_{\Delta T} c(t)dt}{\Delta T}.$$

Our DSS is aimed at minimizing this internal KPI through policy selection.

In order to evaluate the accuracy in selecting the best policy in portfolio of available strategies, we designed a second set of experiments using the same setting (100 days and four snapshots) as that in the previous section. In particular, for each day and snapshot, we first run the DSS for each policy and determine the best one according to the internal KPI; then, we "run" the real ED (i.e., we run its digital twin) in the same setting, and determine the best policy for the real system. In our tests, we assume that policy maker of our DTS always accepts and follows the DSS suggested policy. Finally, we count the number of times in which the resulting strategies coincide, meaning that the DSS was able to identify the best policy for the real ED. Table 4.3 reports the results obtained with different values of $\Delta T = \{1,2,3\}$ hours for both the deterministic and the stochastic versions of the DSS. As may be expected, performances get worse when increasing the value of ΔT. In addition, the results confirm that Approach 3 in the stochastic operating mode provides the best policy prediction, with more than 80% of success for 1 and 2 hours.

	Approach 1		Approach 2		Approach 3	
ΔT	% Det.	% Stoc.	% Det.	% Stoc.	% Det.	% Stoc.
1	73.00	70.00	72.50	76.00	75.00	82.00
2	73.50	71.00	75.50	71.50	77.50	82.00
3	62.00	67.50	69.50	62.50	68.50	73.50

TABLE 4.3: Accuracy of each approach in identifying the best policy.

Improving performance

In this section we present the experiments performed in order to evaluate the capability of the DSS to improve the ED performance. For this aim, we set-up an experimental environment replicating the configuration depicted in Figure 4.2, where we replaced the real ED by its DTS. The system was populated with data of the patients arrived between 01/10/2019 and 15/10/2019; this interval, which was excluded from the previous experiments. We tested the DSS for the whole period of 15 days, thus evaluating the potential cumulative effect of decision. Following the indication provided by the previous experiments, the DSS embedded Approach 3 for the predictor.

Our first order of business is to determine the best setting for parameters ΔT and Δz. To this aim, we tested the DSS in both the deterministic and stochastic configurations, for different values of ΔT and Δz. Tables 4.4 and 4.5 show the results in terms of mean number of patients within the ED during the 15 days. The best value for each Δz is highlighted in yellow, while the best overall in green.

ΔT / Δz	Real	30min	60min	120min	180min	240min
30min	44.36	38.58	37.85	37.24	41.23	42.40
1h	44.36	37.14	37.17	36.20	35.46	34.40
2h	44.36	37.30	38.71	37.15	35.34	35.80
3h	44.36	37.26	38.77	35.80	34.58	35.59
4h	44.36	41.44	42.52	41.68	40.82	38.79

TABLE 4.4: Δz vs ΔT, Det. configuration

ΔT / Δz	Real	30min	1h	2h	3h	4h
30min	44.36	38.20	39.51	37.61	43.86	42.06
1h	44.36	37.44	34.60	37.50	35.35	35.63
2h	44.36	38.37	37.74	35.40	36.44	35.93
3h	44.36	37.73	38.31	36.63	35.23	35.44
4h	44.36	41.47	42.68	42.52	46.35	38.23

TABLE 4.5: Δz vs ΔT, Stoc. configuration

Table 4.4 shows that in general the best results are obtained when $\Delta T \geq \Delta z$ (obtained with the deterministic configuration), although in all cases the DSS is able to reduce the mean number of persons compared to the real ED.

Figure 4.5 plots the number of patients in the real ED and the same figure obtained through the DSS, in the best Deterministic configuration ($\Delta T = 4\text{h}$, $\Delta z = 1\text{h}$), and in the best Stochastic configuration ($\Delta T = 1\text{h}$, $\Delta z = 1\text{h}$). Summarized statistics, also reporting the average Length of Stay (in minutes) and Waiting time (W.T. in minutes), can be found in Table 4.6.

The results confirm that, although internally optimizing the mean number of patients, the DSS also considerably improves over the real ED for what concerns the LoS and W.T. of patients. In Table 4.7 the mean and median LoS of patients are disaggregated by urgency code, showing that introducing the

Indicator	Real	Det.	Stoc.	Det. gain %	Stoc. gain %
C_{avg}	44.36	34.40	34.60	-22.45	-22.00
LoS [min]	322	263	268	-18.29	-16.80
W.T. [min]	146	128	139	-12.68	-4.86

TABLE 4.6: 15 days analysis results

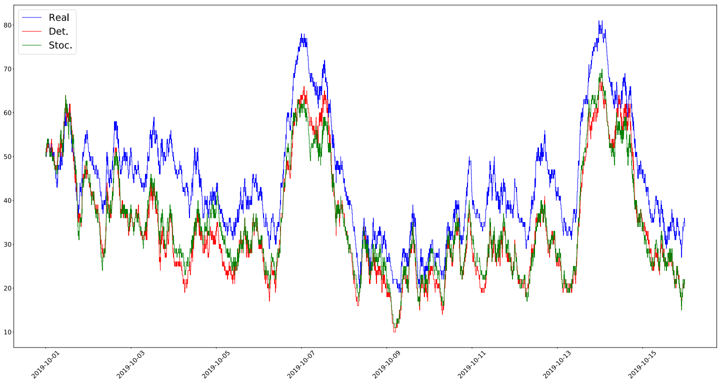


FIGURE 4.5: Comparison between the real trend of the patients present and those obtained with DSS

DSS has a major impact on the low priority patients, without affecting the (already short) LoS of urgent ones. Table 4.8 shows the mean and median Waiting time of patients disaggregated by urgency code and it proves how the DSS has good effect especially on high priority patients.

	Real LoS [min]		Det. LoS [min]		Stoc. LoS [min]	
	mean	median	mean	median	mean	median
Urgency Code						
1	173	144	174	146	176	158
2	261	224	242	203	245	206
3	385	356	290	276	295	288
4	379	280	273	263	290	262

TABLE 4.7: LoS performance in minutes on the basis of the patients' urgency

In order to asses the quality of our results, we evaluate the average number of patients within the ED obtained by applying the (ideal) policy which can move patients through different queues, as many times as needed and with null transfer time (FQR). This figure reduces the C_{avg} from 34.40 to 33.02, showing that only a marginal improvement could be obtained by a redefinition of the hospital layout and logistics.

Finally, we run the DSS for each day of the period individually, for both the Deterministic and Stochastic configurations in their best settings. Table 4.9 shows the mean number of patients within the ED, highlightning the best figure each day. In all but two cases the DSS provides a considerable improvement; in the remaining two days, the DSS results are only slightly

Urgency Code	Real W. T. [min] mean	Real W. T. [min] median	Det. W.T. [min] mean	Det. W.T. [min] median	SM W.T. [min] mean	SM W.T. [min] median
1	9	6	2	1	3	1
2	35	33	22	21	30	28
3	223	197	201	191	218	195
4	312	176	289	200	291	201

TABLE 4.8: Waiting time performance in minutes on the basis of the patients' urgency

worse than the historical ones, thus confirming the robustness of the proposed approach.

Day	Real	Det.	Stoc.	Det. gain %	Stoc. gain %
2019-10-01	50.28	47.34	47.55	-5.85	-5.43
2019-10-02	47.78	43.12	43.40	-9.75	-9.17
2019-10-03	49.78	43.97	45.45	-11.67	-8.70
2019-10-04	43.24	39.82	41.87	-7.91	-3.17
2019-10-05	38.25	34.17	33.70	-10.67	-11.90
2019-10-06	48.24	46.88	48.94	-2.82	1.45
2019-10-07	62.54	62.88	62.99	0.54	0.72
2019-10-08	32.13	27.98	28.77	-12.92	-10.46
2019-10-09	26.28	21.25	22.16	-19.14	-15.68
2019-10-10	33.03	28.17	29.47	-14.71	-10.78
2019-10-11	39.69	30.51	30.47	-23.13	-23.23
2019-10-12	42.58	33.68	33.67	-20.90	-20.93
2019-10-13	52.15	46.25	49.02	-11.31	-6.00
2019-10-14	63.44	63.84	63.60	0.63	0.25
2019-10-15	36.15	33.24	33.63	-8.05	-6.97

TABLE 4.9: Daily analysis results

4.6 Conclusions

In this chapter we described a decision support system to improve the performance of an ED by addressing the serious problem of overcrowding. This is a complex task as it involves a non-univocal definition of the metrics to be considered, and includes stochastic events.

The DSS includes 4 main elements, namely a predictor for patients arrivals, a predictor for the patients pathways, a predictor for activities duration, and a discrete-event simulator. The DSS implements and tests different policies and returns the one which provides the best expected performance. In addition to structured information collected at triage, the predictor for the patients pathways also exploits unstructured information from the nurse's diary, processed through a Natural Language Processing module.

An experimental application of the DSS to a digital twin of a major real ED in norther Italy has demonstrated that an online selection of the best policy makes it possible a relevant reduction of the number of patients within the ED, as well as their Length of Stay.

The presented DSS is ready for being used in the ED of our case study. In addition, as the prediction-simulation modules implement a quite general framework, we expect the adaptation of the DSS to other EDs to require limited effort.

The first area for future development, which is directly linked to an actual use of the tool, concerns the evaluation of the results obtained following a partial rejection of the tool "suggestions" by the decision-maker. This would make it possible to assess the *rejection rate* with which performance would still be improved. Such a parameter would be useful, since in the real case the decision maker might not always accept the tool "advice", due to clinical reasons.

The study on the selection rate of each policy is another interesting area for improvement. This indicator and the organisational patterns in which a given strategy is selected would make it possible to reduce the set of strategies to be tested by excluding those historically less effective in that condition.

Expanding the set of available policies appears to be an other promising direction of future research. This can be obtained by pre-defining additional policies, or, in a more challenging perspective, by allowing the system to discover its own policies through learning. Concerning the first strategy, it might be worth starting with strategies already known in the literature for their effectiveness, such as the *First Consultation Priority Rule* or the *Second Consultation Priority Rule*(see Cildoz et al., 2018 and Cildoz et al., 2019). For the second area, reinforcement learning (see Sutton and Barto, 2018) appears to be the right approach for the construction of an agent, which learns improvement strategies on its own, after the definition of a set of rules.

Chapter 5

Organization of the response to a pandemic emergency: a classifier for COVID-19 patients

In this chapter we present the organization of the response to a pandemic emergency, through the coordination of different companies. We describe the process of taking a patient in charge and introduce an algorithm that gives each patient a state, through which they are directed to the appropriate path. This tool gave a fast automatic way to manage the system and to define the best clinical strategy. Analysis of a real case study on COVID-19 pandemic in Bologna, Italy, showed that the use of such tool helped healthcare professionals during the most difficult moment of the emergency (the first 6 months), when swabs were scarce and information systems were inadequate.

5.1 Introduction

Since December 2019 COVID-19 pandemics has spread all over the world, with a huge impact on healthcare systems and processes. Many different authorities have had to direct efforts to address this emergency, by taking hard decisions, such as the interruption the surgical activities and the cancellation of appointments of outpatients.

Italy has been among the first, and one of the most severely affected countries in the world (La Rosa et al., 2021), especially in the northern area where first cases were recorded. Due to this unfortunate situation, the northern regions of the country were the first to develop systems to cope with the emergency.

This is the case of the Metropolitan Area of Bologna City. Bologna is the capital of Emilia-Romagna Region, one of the regions in which COVID had a greater impact since the beginning (fifth greatest number of cases in Italy; the third during the first 5 months of the emergency), 1.9 Million of cases in October 2022, and the second greatest number of deaths, over 18,000 in October 2022 (according to Dong et al., 2020 and Ritchie et al., 2020). Emilia-Romagna has a population of 4.47 Million people and almost a fifth of it lives in the Metropolitan Area of Bologna City (about 870,000 inhabitants). In Emilia-Romagna seven Local Health Authorities (LHA)s are in charge of providing services to the population. The LHA of Bologna is composed of 6 territorial districts and 9 hospitals. Within the same city there are two other public organization which provide clinical services to the population: IRCCS[1] Azienda Ospedaliero-Universitaria di Bologna Policlinico S.Orsola (AOSP) and IRCCS Istituto Ortopedico Rizzoli (IOR). AOSP provides many different types of services and it is the seat of School of Medicine of the University of Bologna. IOR is a well-known organization dedicated to the treatment of orthopedic diseases. In addition to public companies, in Bologna there are many private hospitals. This huge environment makes Bologna one of the biggest and complex healthcare system in Italy.

From the first of March 2020 Italian government has gradually closed all activities, until the beginning of lockdown (March 9^{th}), in order to face the emergency and avoid overcrowding in hospitals. In particular, people were instructed to stay at home and go to the emergency department (ED) only in severe cases. This fact created two results:

- people did not know what to do if they suspected they were COVID-19 positive;
- sick people waiting at home risked a worsening of their conditions.

Moreover, supplies of swabs for detecting the disease were scarce at the beginning of the pandemic.

In order to face these problems, since the beginning of April 2020, Bologna LHA and other Bologna health companies organized a coordinated response. The goal was to intercept positive patients[2] as soon as possible and then send them to the most appropriate pathways. A key element of this system was an automatic tool to assess patients' pathways and assigning them a *state* denoting theirs features. Through this classification is possible to manage

[1] This is a title granted by the Italian Ministry of Health to biomedical institutions of relevant national interest. An organization with the title of IRCCS has to dedicate a relevant part of its activity to the scientific research.

[2] At the beginning a patients were considered positive based on the symptoms, while later, when supplies of swabs increased, they were considered sick after the test.

thousands of people, without human intervention. This chapter focuses on this classification tool, which is based on an algorithm, and describes the way it supported the process by providing a quick and coordinate response to one of the most difficult LHAs challenge.

5.2 Literature Review

Since the end 2019, when COVID-19 was identified and named (Lai et al., 2020), health systems have tried to find the appropriate way to deal with the emergency, pending the development of a vaccine. Efforts were aimed at two management drivers: clinical and logistics. A comprehensive review of these first attempts is described in Ochani et al. (2021).

Concerning the clinical efforts, the scientific community initially focused its attention on understanding the modes of transmission and the contagiousness of the virus (see Van Doremalen et al., 2020). Subsequently, the method for detecting positive patients became the major concern, with the development of the correct technique for carrying out swabs and serological tests (see To et al., 2020, Agulló et al., 2021 and Sharma et al., 2021). The last challenge concerned how to treat positive patients from a clinical point of view, in terms of clinical tools and drugs (see Bhimraj et al., 2020 and Drożdżal et al., 2021). Scientists from many disciplines (including operations research, management science or Statistics) approached this issue from different points of view.

One of the first problems that were addressed concerns the prediction of the trend of new positive cases, taking into account the restrictions adopted by governments. This problem was approached through two main models: statistical models and machine learning models. Concerning the first approach a remarkable attempt was proposed by ArunKumar et al. (2021), who studied the dynamics of cumulative COVID-19 cases in 16 countries through Auto-Regressive Integrated Moving Average, and Seasonal Auto-Regressive Integrated Moving Average models. Concerning the second approach, that often makes it possible to overcome the necessary assumptions of statistical models, we can find the contribution of Alassafi et al. (2022) who proposed a deep learning model to develop a prediction model for the spread of the COVID-19 outbreak to and throughout Malaysia, Morocco and Saudi Arabia. Another interesting contribution came from Cinaglia and Cannataro (2022), who proposed a new way for analyzing the epidemic trends of COVID-19, based on combined neural networks and Rt estimation.

Another topic of interest are swabs and their management. For example in Garnett et al. (2020), the authors evaluated the utility of different swabs and transport mediums for the molecular detection of SARS-CoV-2. In Benoni et al. (2022), the authors evaluated the cost-effectiveness of different strategies to ascertain COVID-19 recovery in healthcare workers. In Aringhieri et al. (2022), the authors have raised the problem of digital contact tracing; they introduced a new optimization problem, which is the daily problem of collecting swab tests at home in such a way to guarantee a timely testing to people notified by the app to be in contact with a positive case. Other major contributes analyzes the management of swabs from a clinical point of view only.

By analyzing the literature, however, we see that there is a lack of models that propose an integrated organizational response to the emergency. To the best of our knowledge, there are no algorithms that generate the state of patients by starting from an event log, in order to support the system organization.

5.3 Reaction to the COVID-19 emergency of the Local Health Authority of Bologna

As many other cities in northern of Italy, Bologna was strongly affected by COVID from the very first moment. The first official case was recorded in February 29^{th} 2020, and since that date COVID has spread like a wave. Figure 5.1 shows the trend of COVID positive patients, as daily new cases and cumulative number until November 13^{th} 2022 (figures 5.1.a and 5.1.b) and the focus on 2020 (figures 5.1.c and 5.1.d). As shown, the number of cases recorded during the year 2020 is significantly lower than those recorded during 2021 and 2022 (-69% and -94%). This is due to two different factors: SARS-CoV-2[3] new variants have increased transmissibility compared the original one (McLean et al., 2022); at the beginning, the stocks of swabs, to identify COVID patients, were scarce, therefore anyone with symptoms and their contacts were considered positive and isolated.

When the first case was recorded, Bologna organized a response similar to the one for adopted other virological emergencies, through the activation of two different subjects.

[3] Name of the virus causing the COVID-19 disease.

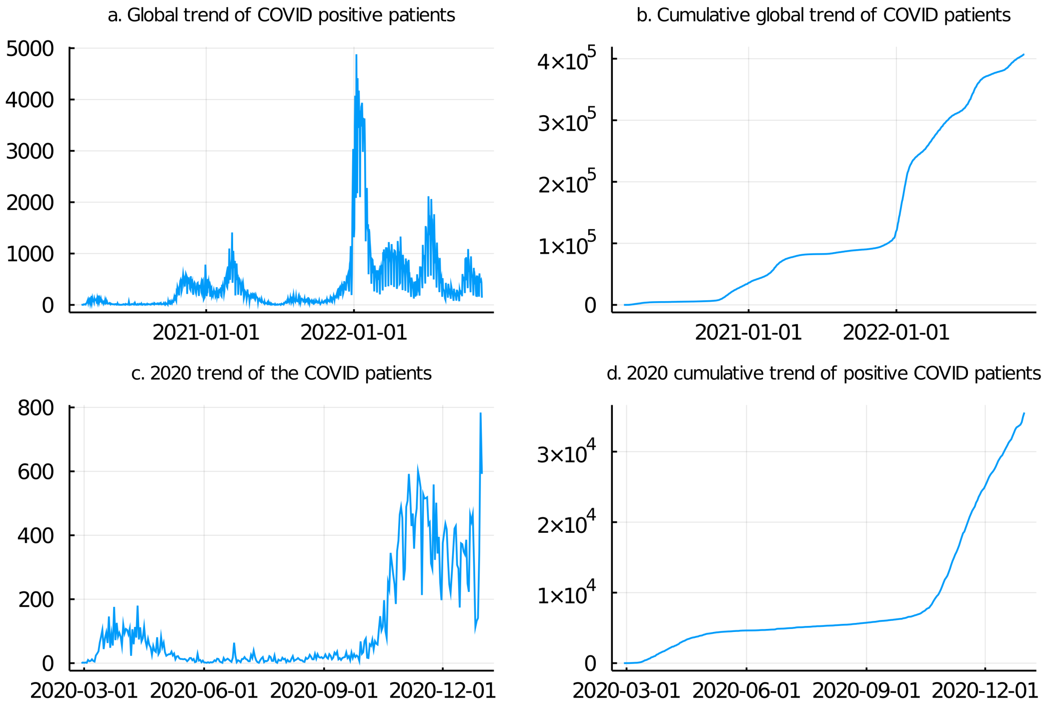


FIGURE 5.1: Trend of COVID positive patients in Bologna

The first one is the Department of Public Health (DPH), that is a part of the Bologna LHA and whose tasks include discovering new cases and monitoring existing ones. From March 2020 to April 2020, this activity was carried out by telephone without the aid of a specialized IT system. Patients identified as possible SARS-CoV-2 positive were isolated at home until clinically healed, just like their contacts who were evaluated and finally isolated. This system soon became very complex to manage, due both to the network of contacts to manage concerning each individual case, and to the large number of places where the cases were identified (ED, Hospitalization, at-home etc...).

The second subject involved from the beginning in the management of COVID was the Department of Infectious Diseases, that is a part of AOSP and which is in charge of managing severely ill patients and carrying out studies on the virus. Due to the high number of hospitalized patients, many other departments (also from other companies) were soon involved to support COVID-19 hospitalized patients.

As previously described, after March 2020 the Italian government gradually closed all activities, until the beginning of lockdown (March 9^{th}), in order to face the emergency and avoid overcrowding in hospitals. People were instructed to stay at home and go to the ED only in severe cases. Through this intervention, the spread of COVID was partially arrested, but at the same time the patients waiting at home when eventually hospitalized had a much more serious clinical condition.

In order to be more effective in facing the emergency, at the beginning of April 2020 Bologna LHA and the other public and private health companies organized a network of response and implemented a *rollback strategy*.

First of all, companies involved general practitioners (GP) as first gate for patients. When a patient had symptoms, they could report it to the GP who reported it to the DPH via an ad-hoc IT system. A clinical algorithm was implemented within this system, in order to evaluate the severity of the patient's symptoms. For more serious patients, GPs could organize an appointment within a special COVID clinic. This organization made it possible to intercept and treat patients before they got worse. Through the same IT system, GPs could report when a patient was clinical healed (without symptoms)

On the other side, companies reorganized the hospital network as a whole, with the departments organized based on the severity of hospitalized patients, from the intensive cares to isolation structures. This organization made it possible to rationalize resources by moving patients to less complex settings once they were better.

This strategy made Bologna more effective in terms of capability to face the emergency, but this new system was much more complex to be manage, due to the large number of subjects, different organization and IT systems involved. Moreover, the number of patients kept growing, making it difficult to follow their clinical paths. To face this issue, we developed a new system which followed and classified patients' pathways. The basic idea was to organize the information into a single event log and assign a state to the patients based on this. The state was used by DPH in order to group patients and assign the to a specific service (e.g. swab agenda).

5.4 Classifier

In this section we describe the Classifier tool we used to support the emergency system. It was developed so as to organize the information in a such way that is accessible, reliable and useful to take decisions. This was an hard task, due to two reasons:

- Sparse patients' information through different organizations' DBs;
- Information systems not designed for this type of management.

These issues were overcome through close collaboration with IT departments, who helped us select and structure the information needed.

5.4.1 State variables

Considering the process and the environment described so far, we decided to split whole patients' state into three different variables, in order to simplify patients management and avoiding a huge number of combination:

- Logistic variable;
- Virological variable;
- Serological variable.

The Logistic variable combines information about patient's position and clinical picture, in order to know where they are and how they are. We designed seven different values that this variable can assume:

At home with symptoms (LS1) identifies a patient who is not within an hospital (they can be in a nursing home) now, but they were reported as a symptomatic case;

At home without symptoms (LS2) identifies a patient who is not within an hospital (they can be in a nursing home) now, and they were not reported as a symptomatic case;

At home clinically healed (LS3) identifies a patient who is not within an hospital (they can be in a nursing home) now, and they healed from her/his symptoms (reported by the GP or DPH);

Within an ED (LS4) identifies a patient who is within an ED now;

Hospitalized (LS5) identifies a patient who is within an hospital now;

Waiting for a medical appointment (LS6) identifies a patient awaiting scheduled access to a COVID clinic;

Closed (LS7) identifies a patient who passed away.

Patients in LS4, LS5 and LS6 states implicitly have clinical symptoms.

The Virological variable describes patients' state from the point of view of the swabs performed. We designed six different values that this variable can assume:

Negative for SARS-CoV-2 (VS1) identifies a patient who had at least a swab performed and who has never had a positive swab;

Positive for SARS-CoV-2 (VS2) identifies a patient whose last swab is positive;

Healing (VS3) identifies a previous positive patient whose last swab is negative;

Healed (VS4) identifies a previous positive patient whose last two swabs are negative;

Deceased (VS5) identifies a patient who passed away;

Unknown (VS6) identifies a patient who has not had any swabs yet.

The Serological variable describes patients' state from the point of view of the serological tests performed. The serological tests are used to identify the presence of antibodies fighting a virus, in this case the SARS-CoV-2 virus. This test identify the presence of two types of antibodies: IgG and IgM. IgG are antibodies produced days after infection and provide the memory and immune defence. IgM are the first antibodies to be produced in response to an infection; their presence indicates a recently developed infection.

We designed four different values that this variable can assume:

Negative (SS1) identifies a patient who has at least a serological performed and who has never had a positive serological;

Positive for IgG (SS2) identifies a patient whose last serological is positive for IgG;

Positive for IgM (SS3) identifies a patient whose last serological is positive for IgM;

Unknown (SS4) identifies a patient who has not had any serologicals yet.

It is possible that a serological test could be positive for both IgG and IgM. In this case the patient is considered as "Positive for IgM", because it means that there is an ongoing infection.

5.4.2 Logs

Within a clinical pathways, particularly the COVID ones, many different physicians are involved (e.g. GPs, specialists of different organizations) through different hospitals. This particular organization creates a lot of data spread across different databases. We have studied all information and selected only the information that is relevant to describe a patient's state.

All data selected were structured in an event log, that is the basis of the algorithm described in the next section. The events considered are listed below, grouped by data source.

DPH source

DPH database contains information about patients with a possible COVID-19 infection. This database is managed by physicians of the DPH. In this source we can find these events:

- Patient reporting (**e1**) - event triggered when a patient is reported as a possible COVID-19 infected to LHA (also a GP could report a patient);
- Reporting of symptoms (**e2**) - event triggered when the patient at home reports suffering from symptoms
- Reporting of no symptoms (**e3**) - event triggered when the patient at home reports that they do not suffer from symptoms
- Clinically healing (**e4**) - event triggered when the MD/GP reports that the patient at home is clinically healed;

Patients Registry

Patients Registry contains all patient's personal data. This database is useful to know if a patient is passed away outside an hospital:

- Patient deceased (**e5**) - event triggered when a patient dies outside an hospital.

ED sources[4]

ED database contains information about ED access, diagnosis and how the access ends. We consider only access with diagnosis related to SARS-CoV-2. For this database we can have these events:

- Beginning of ED access (**e6**) - event triggered when a patient enters the ED;
- Ending of ED access with hospitalization (**e7**) - event triggered when a patient is discharged and needs hospitalization;

[4]This repository changes based on the organization being considered, but the events are the same.

- Ending of ED access without hospitalization (**e8**) - event triggered when a patient is discharged and does not need hospitalization;
- Ending of ED access with death (**e9**) - event triggered when a patient dies in the ED.

Patients hospitalization sources[4]
Patients hospitalization database contains information about the beginning and the ending of hospitalization, the hospitalization setting (e.g. "COVID high intensity") and how the hospitalization ends. For this database we can have these events:

- Beginning of Hospitalization (**e10**) - event triggered when a patient is hospitalized;
- Changing of Hospitalization (**e11**) - event triggered when a patient is discharged, but starts another hospitalization within another hospital;
- Ending of Hospitalization with death (**e12**) - event triggered when a patient dies during the hospitalization;
- Ending of Hospitalization (**e13**) - event triggered when a patient is discharged.

Laboratory Information System (LIS)
LIS contains information about swabs and serological tests. We consider only the positive and negative ones:

- Positive swab (**e14**) - event triggered when a patient receives a positive swab;
- Negative swab (**e15**) - event triggered when a patient receives a negative swab;
- IgG-Positive Serological test (**e16**) - event triggered when a patient receives a IgG-Positive Serological test;
- IgG-Negative Serological test (**e17**) - event triggered when a patient receives a IgG-Negative Serological test;
- IgM-Positive Serological test (**e18**) - event triggered when a patient receives a IgM-Positive Serological test;

- IgM-Negative Serological test (**e19**) - event triggered when a patient receives a IgM-Negative Serological test.

GP sources
GP database contains information about patient healing reported by GP and medical appointment scheduled by GP:

- Clinically healing (**e4**) - event triggered when the MD/GP reports that the patient at home is clinically healed (repeated);
- Scheduling medical appointment (**e20**) - event triggered when a GP arranges a scheduled medical appointment for a patient;

Others sources
Information spread in different databases. For instance a medical appointment can be provided by different users depending on priority level and each of them has a different database.

- Providing medical appointment (**e21**) - event triggered when a medical appointment is provided.

The event log we obtained has a structure based on Van der Aalst (2016) studies, so this and the algorithm described in the next section can be applied with a small number of modifications to many other systems[5]. Within the event log each row represents a specific event that happened to a specific patient and it has the following attributes:

- Case_ID, which is a patient identifier;
- Event, which is the type of event that occurred to that patient, among those described so far;
- Timestamp, which is when the event occurred;
- Place, which is when the event occurred;
- User[6] who performed the activity.

[5]It is also easily add other event type, in order to customize the system.
[6]Not always filled.

5.4.3 Classification algorithm

In this part we show the pseudo-code of the classification algorithm that process the logs and defines the value of the state variables. The algorithm is divided into 3 parts, one for each variable type. In each of the three parts a patient starts with the last known state[7] and it changes based on the new recorded event. Each state change is logged in order to evaluate the system performance, even if two events occur in close moments.

Algorithm 2 Algorithm to find the value of the patient's logistic variable

Input:

- Event log described in subsection 5.4.2 as *LOG*;
- Patient's last Logistic variable as *state*;
- Logistic variables described in subsection 5.4.1.

Algorithm 2 is based on the analysis of the patient's last state and the last event that occurs. For example, a patient who is identified as being within an ED (LS4) and for whom the new event records a discharge home (e8), is transitioned to the state of "at home with symptoms" (LS1). The choice of LS1 state is due to the fact that it is likely that the patient is not yet clinically healed after the ED pathway, but their symptoms are not severe enough to be hospitalized.

On the other hand an hospitalized patient (LS5), who is discharged (e13), changes the state into "at home without symptoms" (LS2). Indeed, a hospitalized patient is only discharged when they are completely healed of symptoms.

Algorithm 3 Algorithm to find the value of the patient's virological variable

Input:

- Event log described in subsection 5.4.2 as *LOG*;
- Patient's last Virological variable as *state*;
- Virological variables described in subsection 5.4.1.

[7] For new patients, a combination of LS2-VS6-SS4 is assigned as default states.

Algorithm 2

```
for event ∈ LOG do
    if (state != LS7) then
        if (event == e5) or (event == e9) or (event == e12) then
            state := LS7;
        else if ((event == e1) or (event == e14) or (event == e15) or (event == e16) or (event == e17) or (event == e18) or (event == e19)) and (state == void) then
            state := LS2;
        else if ((event == e7) or (event == e10) or (event == e11)) and (state != LS5) then
            state := LS5;
        else if (event == e6) and (state != LS4) then
            state := LS4;
        else if (event == e2) and ((state == void) or (state == LS2) or (state == LS3)) then
            state := LS1;
        else if (event == e3) and ((state == void) or (state == LS1)) then
            state := LS2;
        else if (event == e4) and ((state == void) or (state == LS1) or (state == LS2)) then
            state := LS3;
        else if (event == e20) and (state != LS4) and (state != LS5) then
            state := LS6;
        else if (event == e21) and (state == LS6) then
            state := LS1;
        else if (event == e13) then
            state := LS2;
        else if (event == e8) then and ((state == LS4)
            state := LS1;
```

Algorithm 3

```
for event ∈ LOG do
    if (event == e14) and (state != VS2) then
        state := VS2;
    else if (event == e15) and (state == VS2) then
        state := VS3;
    else if (event == e15) and (state == VS3) then
        state := VS4;
    else if (event == e15) and ((state == VS6)or(state == void)) then
        state := VS1;
    else if (event == e5) or (event == e9) or (event == e12) then
        state := VS5;
    else if state == void then
        state := VS6;
```

Algorithm 3 works similarly to Algorithm 2. For example, a patient in the positive state (VS2), who receives a negative swab (e15) switches to a healing state (VS3). Then, with a new negative swab (e15), they will pass into healed state (VS4).

Algorithm 4 Algorithm to find the value of the patient's serological variable

Require:

- Event log described in subsection 5.4.2 as *LOG*;
- Patient's last serological variable as *state*;
- Serological variables described in subsection 5.4.1.

Algorithm 4

```
for event ∈ LOG do
    if (event == e18) and (state != SS3) then
        state := SS3;
    else if (event == e16) and (state != SS2) then
        state := SS2;
    else if ((event == e17) or (event == e19)) and (state != SS1) then
        state := SS1;
    else if state == void then
        state := SS4;
```

Algorithm 4 analyzes serological test events and states. For example a patient who is not positive for IgM (SS3) and they receive an IgM-positive serological test (e18) and then they assume the state of positive for IgM (SS3).

5.4.4 Architecture and performance

We implemented the algorithm in the Python 3 language[8] (Van Rossum and Drake, 2009), and run it on a server with an AMD Ryzen 7 2700C 8 Core processor, 64GB RAM under the Ubuntu Linux operating system. The processing time has to be divided into two different moment:

1. First running
2. Daily running

The first one concerns the events processing when the tool was adopted the first time (April 14^{th} 2020), with a huge number of patients and events. In particular the event log had the size of 100,000 events recorded for 31,000 patients and it was processed in 5 hours.

After we have created the first list of states, we have decided to run the algorithm every day during the night[9], in order to process the events recorded during the day (mean number of 1,900 events). In this case the process lasted about 20 minutes.

5.5 Impact

In this section, we present how the tool was used to support the organisation response to the COVID-19 emergency and how it improved performance.

5.5.1 Patient management

As previously described, based on the combination of the three variables, patients were referred to the appropriate pathway. The pathway organization changed during pandemic months in order to adapt it to the latest clinical indications. In Table 5.1 we show the pathways and their variables that were used for the longest time during 2020.

[8] This is not the best choice for this type of algorithm, but the technical conditions in terms of data availability make it the simplest.

[9] We decided to run the algorithm once a day because it was compliant with our goal, but in view of the low processing time, this number could have been higher.

As can be seen, pathways are very different, but they can divided into three different groups, based on their goals:

1. Organization of the serological agenda (1-2)

2. Organization of the swabs agenda (3-4-5)

3. Patient monitoring (6-7)

The proper definition of the first group was of paramount importance during the first four months of the emergency, when swabs were scarce and LHA needed to filter patients before organizing swab appointments.

The second group concerns the organization of the swabs agenda, for suspected COVID patients. During the first part of the emergency, not all patients were swabbed, but if they had symptoms and they were not at risk of aggravation, they were simply isolated. In case of need, an appointment was arranged in the COVID clinics.

The third group concerns the patients monitoring, in order to end the isolation or check, through a telephone interview, at-home patients.

In addition to the listed pathways, the use of the tool made possible to keep track of patients who would otherwise had been lost (e.g. homeless).

Logistic variable	Virological variable	Serological variable	Pathway
LS1 or LS2 after an hospitalization	VS6	SS4	Schedule an appointment for a serological test after 7 day
LS1 or LS2 not-after an hospitalization	VS6	SS4	Schedule an appointment for a serological test after 14 days
LS1 or LS2	VS1 or VS6	SS2 or SS3	Arranging an Schedule an appointment for a swab as soon as possible
LS3	VS2	Any	Schedule an appointment for a swab after 14 days
LS3	VS3	Any	Schedule an appointment for a swab after 48 hours
LS3	VS4	Any	Report to DPH for notify the end of isolation
LS1 (not-reported by GP)	VS6	VS6	Report to the GP to follow up the patient's healing

TABLE 5.1: Patients' pathways based on state variables

5.5.2 System performance

Concerning the performance of the system, it can be analyzed from a qualitative and a quantitative point of view.

First, the level of service offered to the patient must be considered. Through centralized data and state management, it is possible to know with certainty where a patient is at any time. This prevents a patient from being sent to one pathway while being taken care of in another. For example, in the first part of the pandemic it happened that patients, who had been hospitalized, were contacted for follow-up by DPH, who did not know their state.

Another improved aspect was the efficiency of the offered service. Before the development of the classifier, grouping patients according to the appropriate pathways could take hours (sometimes even days), since they were performed by hand by an operator. This led to a very slow management of cases, with the risk of non-interception of possible positives and aggravation of patients. Through the classifier, every day the patients were grouped automatically and each operator could concentrate on the group of their pertinence, without any preliminary operation.

The last qualitative measure concerns the ability of the system to adapt to changes. During the months of 2020 (and beyond), the guidelines regarding COVID were changed. This led to the system having to reorient itself several times. Through the use of the classifier, these changes were more effectively implemented.

From a quantitative point of view, it is difficult to measure how much the classifier provided support. This is due to two aspects: the guidelines on how to treat COVID-19 patients have changed; the available resources have changed over time. Nevertheless, we do provide some measures as a proxy for how the system performed before and after (April 14^{th} 2020, until December 31^{st} 2020) the classifier was adopted. Table 5.2 shows statistics about the time in days between interception of patients and first swab/serological[10]. As can be seen, after the introduction of the classifier, the time to the first

Classifier adoption	Mean	Median	Dv_Std
Before classifier	15.8	8.5	83.3
After classifier	2.9	1.4	4.8

TABLE 5.2: Time in days between interception of patients and first swab or serological

[10]Patients with the first event being hospitalization were excluded, as the swab was carried out for them as a screening

swab/serological has dropped significantly. Furthermore, a smaller standard deviation indicates more standardized and structured pathways.

An other interesting indicator is the number of events (see section 5.4.2) within patients' pathways[11] (Table 5.3). As it can be seen, although different

Classifier adoption	Mean	Median	Dv_Std
Before classifier	5.0	4.0	3.7
After classifier	3.3	2.0	3.0

TABLE 5.3: Number of events within patients' pathways

types of paths involving many different actors and events were introduced in April, an increase in events was avoided through state management.

5.6 Conclusions

In this chapter we have described the initial difficulties in dealing with a global pandemic, in a complex context such as the healthcare one. The uncertainties derive from the lack of knowledge of effective protocols both from a clinical and organizational point of view. This complexity increases within a system with different companies which use different IT systems.

It is in such a context that the healthcare system of the city of Bologna found itself operating in February 2020, when the COVID emergency hit the world. To overcome these difficulties, we have introduced a classifier which, starting from the patient's pathway (organized as event log), could deduce a unique state. Through the state, physicians could assign the patients to the correct clinical and organizational solution. Such a tool was only possible through a great deal of preliminary information centralization and structuring work.

This tool has allowed the LHA of Bologna and the other healthcare companies to adapt effectively and efficiently to the emergency in short time. Furthermore, after December 2020 with the arrival of vaccines and the introduction of a new IT system for DPH, the algorithm has become the core of COVID patient management.

[11] A patient's pathway consider each COVID event related to that patient.

Chapter 6

A decision support system for scheduling a vaccination campaign during a pandemic emergency: the COVID-19 case

[1]

In this chapter we present the organization and scheduling of a vaccination campaign during a pandemic emergency. We describe the decision process and introduce an optimization model, which showed a powerful multi-scenario tool for scheduling a campaign in detail within a dynamic and uncertain context. The solution of the model gave the decision maker the possibility to test different settings and have a configurable solution within few seconds, compared with the man-days of effort that would have required a manual schedule. Analysis of a real case study on COVID-19 vaccination campaign in northern Italy showed that the use of such optimized solution allowed to cover the target population within a much shorter time interval, compared to a manual approach.

6.1 Introduction

Scheduling a vaccination campaign during a pandemic emergency is a hard organizational task. We have a limited amount of vaccine shots and a forecast of availability in the next weeks, which is extremely uncertain. As in the case of the COVID-19, vaccines can be of several types, each one with its own prescribed delay between consecutive shots, and prescribed target

[1]The results of this chapter appears in: C.Fabbri, P.Ghedini, M.Leonessi, E.Malaguti, P.Tubertini, "A decision support system for scheduling a vaccination campaign during a pandemic emergency: The COVID-19 case", *Computers & Industrial Engineering*, 177, 2023

population based on age or clinical status. Hence, among the several groups of individuals which are candidate to receive the vaccine, one has to decide which individuals and when inoculate the vaccine, and has to organize and schedule the associated operations.

The COVID-19 vaccination campaign has been a task of great complexity for national and LHAs since December 2020. As for the Italian context, decisions relative to the planning and implementation of the COVID-19 vaccination campaign are taken at three main decision levels: the national board, composed by the Ministry of Health and the COVID commissioner; the regional coordination center; and the Local Coordination Center (LCC). The national board was a reference point for regional representatives in charge of monitoring the progress of procurement, vaccination and surveillance and responsible for strategic decision in terms of:

- volume of COVID-19 vaccines purchased from pharmaceutical companies (e.g. Comirnaty, Spikevax, Vaxzevria, Janssen);
- delivery schedule of purchased doses;
- guidelines for the vaccination campaign, including target age groups to be vaccinated, potential starting date for each age group or target population campaign (e.g. extremely vulnerable people), compatibility between age group or target population campaign and COVID-19 vaccines (e.g. Vaxzevria incompatible for extremely vulnerable people);
- schedule and case mix of supplies distributed among the twenty-one regional health authorities;
- weekly and daily target vaccination volumes for each regional health authority.

The regional coordination centers were the reference entity for the regional vaccination points in charge of monitoring the progress of procurement, vaccination and surveillance within the region. The regional center operated as an organizational bridge between national instructions and guidelines and local operational implementation. At this level decisions were taken in terms of:

- regional starting date for each age group campaign according to national guidelines;
- weekly and daily target vaccination volumes for each LHA;

- booking system to be used for each age group or target population campaign.

LCCs were in charge of the operational coordination in terms of work teams, ordering and making vaccines and other necessary products available, and highlighting any operational problems. LHA oversaw the operational implementation of COVID-19 vaccination campaign. At this level decisions were taken in terms of:

- number of vaccination points to be opened and geographical distribution in order to maximize the coverage of each catchment area;
- characterization of vaccination facilities per vaccination campaign (e.g. mass vaccination facilities for younger people, hospital clinics for extremely vulnerable people, proximity clinics for elderly people);
- maximum volume of daily vaccinations planned for each clinic based on layout capacity, workforce availability (medical, nursing and administrative staff) and stock availability of COVID-19 vaccines;
- configuration of booking diaries for each vaccination campaign;
- implementation of proactive invitation campaigns for extremely vulnerable people.

This chapter focuses in particular on the decisions taken at the third level, concerning the LCC's operations, for which a decision support tool based on a Mixed-Integer Linear Programming (MILP) model is developed.

6.2 Literature Review

The development of vaccines was the most effective tool for controlling the COVID-19 pandemic emergency. However, efficiently exploiting this opportunity created new management challenges, calling for quantitative approaches.

The first challenge faced by the scientific community concerned the prioritization of people to access vaccination. This decision is important in particular at the beginning of a campaign with a new vaccine, due to the low-capacity production. A review of modeling methods to optimize the allocation strategies based on different utility measures has been conducted by K. Liu and Lou (2022). A relevant attempt has been performed by Shim (2021) who proposed a model that aims to obtain the best vaccine allocation in order to optimize three alternatives objectives such as infections, deaths and life expectation. A data-driven model of COVID-19 transmission to deal with the vaccination prioritization problem has been proposed for the Chinese context by Han et al. (2021). The problems of choosing the best candidates for the vaccination not only on the basis of age has been considered by Książek et al. (2022), who propose two MILPs , one focusing on the social groups and the second one on territorial units.

Another relevant topic involved the vaccine supply chain (VSC). Georgiadis and Georgiadis (2021) addressed the problem of minimizing the total cost of VSC trough a MILP, by simultaneously addressing the planning of vaccine supply chains, and the planning of daily vaccinations in the available vaccination centers. Ibrahim et al. (2022) introduced a multi-product MILP vaccine supply chain model for supporting planning, distribution, and administration of different vaccines, having different conservation and distribution requirements. Tavana et al. (2021) developed a mathematical programming approach for fair distribution in developing countries, taking into account the different vaccine storage conditions. A specific item of VSC is the facility location problem. Soria-Arguello et al. (2021) focused on the cross-dock warehouse selection in Mexico in order to minimize costs, while Tang et al. (2022) realized a bi-objective MILP to choose vaccination point, while considering both the economic and service quality criteria.

Bertsimas et al. (2021) proposed an integrated method that considers both prioritization and supply chain through a novel data-driven approach.

Moving to operational issues, Zhang et al. (2022) addressed an overall optimization of the appointments organization, while considering four different objectives: fixed costs for opening a vaccination site, total travel distance of vaccine recipients, total appointment rejection cost, and total tardiness cost. Small instances were addressed through a MILP, while a matheuristics algorithm was used for larger ones.

To the best of the authors knowledge, in this manuscript we present the first MILP model for scheduling a vaccination campaign, while minimizing the vaccination delay, that has been applied in a real setting and for which a solution optimized through a mathematical programming approach has been implemented in a large campaign.

6.3 COVID-19 vaccination campaign for the Local Health Authority of Bologna

In the following sections we describe a MILP model used to support the operational programming of the COVID-19 vaccination campaign for the population of Bologna LHA. This LHA is responsible of health management and health services provision in Bologna, a major city in northern Italy. As already described, this is one of the largest health agencies in Italy by size, its territory includes 46 municipalities for a population of over 870,000 inhabitants, and it is divided into 6 territorial districts, each one headed by a director, including the urban area of the city, a plain exurban area, and a mountainous area. The activities of the coordination center of the Bologna LHA for the COVID-19 vaccination campaign were structured according to a planning approach based on the sequential activation of vaccination sub-campaigns. Each sub-campaign was characterized by a priority grade and the definition of target population (age group or pathology-related) decided at a nation level.

The sub-campaigns were based on a appointment scheduling and booking paradigm where appointment slots for vaccination were made available to the population for self booking through multiple channels (online, telephone, de-visu at booking desks or pharmacies). In addition, a compatibility matrix was given for each campaign, defining the kind of vaccines that could be used for the target population. A summary of sub-campaign data for Bologna LHA is reported in Table 6.1.

Sub-campaign	Population	Comirnaty	Vaxzevria	Spikevax	Janssen
OVER 80	63714	X			
75-79	35278	X	X		
OVER 70 + VULN.	69126	X	X		
OVER 60	34962	X	X		X
OVER 65	24548	X	X		X
FRAGILE	15752	X			
OVER 55	39306	X		X	
OVER 50	30426	X		X	
OVER 40	68756	X		X	
OVER 18	78586	X		X	
OVER 12	127506	X			

TABLE 6.1: Compatibility matrix for each population group and vaccines.

From an operational point of view, the LCC had therefore the responsibility to define the best distribution of the openings on the agenda for each sub-campaign, while taking into consideration the territorial proximity of the vaccination centers to the population, so as to vaccinate the largest possible number of citizens in the shortest time. In defining the appointment slots to be made available, it was necessary to take into account a series of operational constraints: (i) availability of vaccine doses in stock, (ii) maximum capacity of available vaccination sites and (iii) availability of medical, assistance and administrative staff for the management of the vaccination points.

The programming flow of the vaccination sub-campaigns was structured as follows. The regional coordination center, following the input of the national board, outlined the guidelines of the sub-campaign (population group, starting date and booking channels), and communicated them to the LCC that was responsible for operational implementation. The operational implementation of campaign programming for LHA of Bologna can be distinguished in two phases, an initial one (phase 1) in which no support tool was available, and a second one (phase 2) in which a mathematical programming model acted as a pivot in the decision-making process.

The phase 1 decision-making process was essentially based on the definition of the weekly vaccine doses budget available for each of the 6 territorial districts. The vaccine doses budget available was defined by the LHA of Bologna Operations Research team based on the forecasts of deliveries. The task of each District Director was then to identify and verify the availability of the territorial vaccination centers and coordinate the representatives of the other involved departments to check the availability of staff for shifts coverage. Once a proposal was formulated, each District Director communicated to the Operations Research team the appointment time slots to be opened which were checked to assess the consistency with respect to the stock availability. The validated appointment slots plan was then sent in configuration to guarantee bookability through the indicated booking channels. The duration of this decision-making process could take from 3 to 5 days with a time requirement of at least 6 hours for the development of a proposal of appointment slots to be opened being consistent with the availability of the vaccination centers for the second doses. A summary of the phase 1 decision-making process is describe in Figure 6.1.

The phase 2 decision-making process was activated with the "OVER 70 + vulnerable people" vaccination sub-campaign, and envisaged a substantial revolution thanks to the availability of a decision support tool based on a MILP model. The tool allowed the decision maker to schedule a vaccination campaign in few minutes. The campaign specifications were translated into quantitative parameters by the Operations Research team that was responsible for feeding the optimization model with data of the demand to be met (in terms of vaccination coverage), the available vaccination sites, their maximum vaccination capacity per shift, the plan of first and second doses already foreseen by previous sub-campaigns and the forecast of the weekly delivery volumes of vaccine. The hypothesis of vaccination sub-campaign schedule of each district, as computed by the model, was sent to each District Director who was therefore relieved from the tasks of elaborating a proposal and verifying compatibility. This way, the entire process of planning and configuring appointment slots could be completed in two days with a substantial risk error reduction. A summary of the phase 2 decision-making process is describe in Figure 6.2. The MILP model described in the next section was defined and used to support this planning activity.

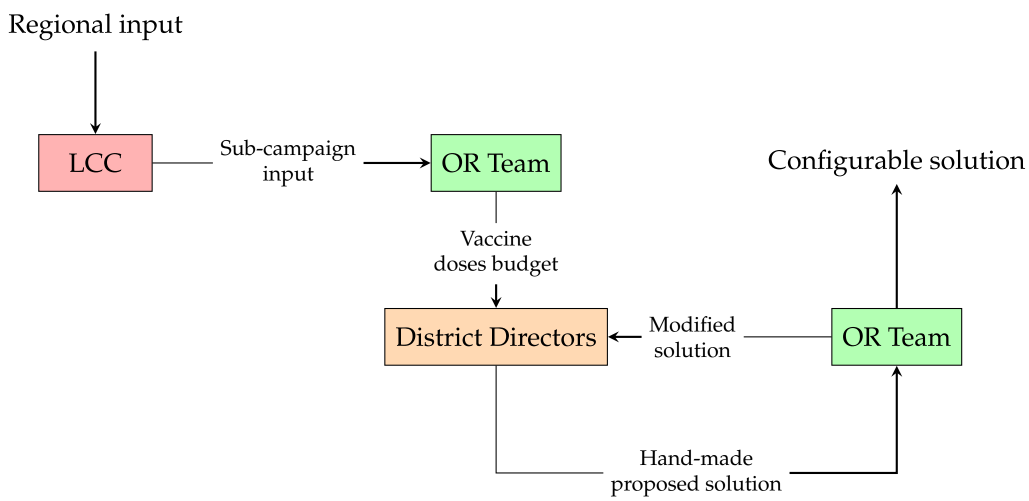


FIGURE 6.1: Decision flowchart phase 1

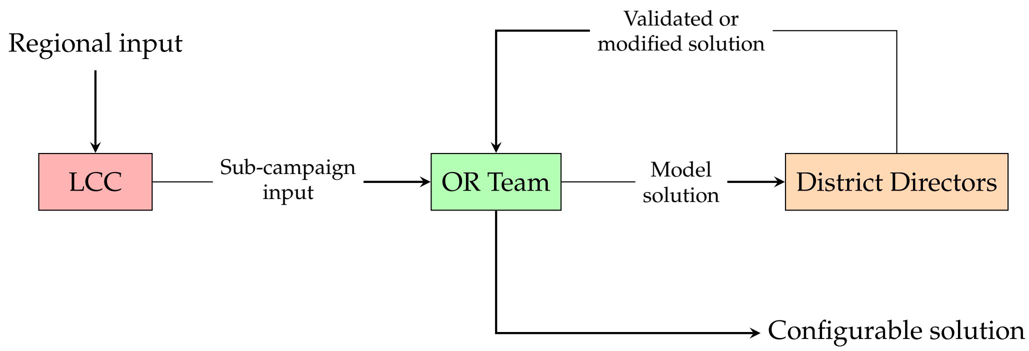


FIGURE 6.2: Decision flowchart phase 2

6.4 Problem formulation and MILP model

A MILP model was developed in order to support the COVID-19 vaccination campaign, by identifying the calendar of opening agendas in line with stock availability and forecast, type of vaccine that can be used by target age group, sub-district demand, divided according to proximity to vaccination hubs and desired coverage rate. The objective of the model is to cover the whole demand in the shortest time, by minimizing a suited function of delay. More in detail, we are given a set S of facilities where the vaccine can be administered, a set T of possible time slots, a set of districts D, each one with an associated demand Q_d, a set V of vaccines and a set Z defining the possible vaccine mixes[2] to be offered at a slot. We plan over a set $G = 1, \ldots, n$ of days; for each slot $t \in T$, we denote by g_t the corresponding day; for each day g, we denote by $T_g \subset T$ the time slots of day g.

Parameters

$C_{st} =$ capacity of facility s at slot t
$Q_d =$ demand for district $d \in D$
$r_{vz} =$ share of vaccine v in mix z
$arr_{vg} =$ delivery of vaccine v for day g
$a_{v1} =$ initial stock of vaccine v
$\Delta_v =$ second shot delay for vaccine v

Decision variables

$$y_{stz} = \begin{cases} 1 & \text{if slot } t \text{ at facility } s \text{ is activated with a vaccine mix } z \\ 0 & \text{otherwise} \end{cases}$$

$x_{stz} =$ number of first shots of vaccine mix z allocated on facility s at slot t

$a_{vg} =$ number of shots of vaccine v available at day g

[2] A vaccine mix refers to a set of vaccines, and their percentage, provided together during a specific opening. For example, when Vaxzevria was used, Comirnaty was also provided (in varying percentages depending on the campaign), for people with specific diseases.

Let $f(g_t)$ be a generic non-decreasing function penalizing the elapsed time associated with day g_t. A MILP model for scheduling the vaccine campaign reads

$$min \sum_{s\in S}\sum_{t\in T}\sum_{z\in Z} f(g_t)x_{stz} \tag{F.O}$$

s.t

$$\sum_{z\in Z} y_{stz} \leq 1 \quad \forall s \in S, \forall t \in T \tag{6.1}$$

$$y_{stz} + y_{s\tau\zeta} \leq 1 \qquad \text{(6.1 bis)}$$
$$\forall s \in S, \forall((z,t),(\zeta,\tau)) \in E$$

$$x_{stz} \leq C_{st}y_{stz} \quad \forall s \in S, \forall z \in Z, \forall t \in T \tag{6.2}$$

$$\sum_{\zeta\in Z}\sum_{v\in\zeta}\sum_{\tau:g_\tau+\Delta_v=g_t} x_{s\tau\zeta}r_{v\zeta} \leq C_{st} - \sum_{z\in Z} x_{stz} \qquad \text{(6.2 bis)}$$
$$\forall s \in S, \forall t \in T$$

$$\sum_{s\in S_d}\sum_{t\in T}\sum_{z\in Z} x_{stz} \geq D_d \quad \forall d \in D \tag{6.3}$$

$$a_{vg} + arr_{vg-1} + \qquad (6.4)$$
$$-\sum_{s\in S}\sum_{t\in T_g}\sum_{z\in Z} x_{stz}r_{vz} - \sum_{s\in S}\sum_{t\in T_{(g-\Delta_v)}}\sum_{z\in Z} x_{stz}r_{vz} = a_{vg+1}$$
$$\forall g \in G \setminus \{1,n\}, \forall v \in V$$

$$x_{stz} \in \mathbf{Z}^+ \quad \forall s \in S, \forall g \in G, \forall t \in T, \forall z \in Z \tag{6.5}$$

$$y_{stz} \in \{0,1\} \quad \forall s \in S, \forall g \in G, \forall t \in T, \forall z \in Z \tag{6.6}$$

$$a_{vg} \in \mathbf{Z}^+ \quad \forall v \in V, \forall g \in G \tag{6.7}$$

The objective function minimizes the cumulative elapsed time of first shots, weighted by the penalty function $f(g_t)$. Constraints (6.1) impose that each slot t in a facility is assigned to no more than one vaccine-mix z, while constraints (6.1 bis) define incompatibility constraints among vaccine-mix and slots at the same facilities, where E is the list of such incompatible pairs.

These constraints are imposed for slots taking place in the same day, in order to avoid possible confusion. Inequalities (6.2) impose capacity constraints for the first shots and link the binary slot activation variables with the integer x variables. Similarly, (6.2 bis) impose capacity constraints, taking into account that the capacity available at slot t is reduced by the number of second shots for a vaccine v for which first shots were dispensed Δ_v days in advance. Constraints (6.3) impose that the demand of each district d is completely satisfied through slots allocated to facilities within the district. Constraints (6.4) define the stock of vaccine at the end of each day. In particular, the LHS of these constraints sums-up the current stock, the deliveries, and subtracts quantities related with first and second shots, which equals the stock at the next day, appearing in the RHS. Finally, (6.5), (6.6) and (6.7) define the domain of the variables.

6.5 Algorithm

The model presented in Section 6.4 was used to organize the vaccination campaign, starting from the "*OVER 70 + vulnerable people*" sub-campaign. Therefore, a basis for direct evaluation of the performance improvement compared with a manual solution is not available. Nevertheless, we have developed a simple heuristic algorithm that mimics the strategy adopted for designing the vaccination campaigns before the model was developed, and provides an alternative realistic scenario for the vaccination campaigns organization.

The algorithm considers the same parameters as the MILP model, and computes a feasible solution, by following an iterative approach similar to the one used by experts. The objective is the same as for the MILP model, i.e., assigning all the demand in the shortest time while satisfying the constraints described in the previous section.

The pseudo-code in Algorithm 5 describes the optimization process. Starting from the district with largest (yet unsatisfied) demand, the algorithm tries to assign a specific number of vaccine shots (the remaining demand or a multiple of the facility capacity) in the first available time slot. This requires to check the facility capacity and the dose availability at the slot and at the time for the second shot. When an assignment is performed, the demand, and capacities are updated and a new district is considered. The algorithm iterates until the demand of all districts is satisfied.

Algorithm 5

Input: demands D_d, $\forall d \in D$; ordered sets: facilities S, time slots T, vaccine-mix Z;
while $\exists d \in D : D_d > 0$ **do**
  pick district d with largest demand D_d
  $PACK \leftarrow []$
  if $D_d \leq 100$ **then**
    $PACK \leftarrow D_d$
  else
    $PACK \leftarrow [100, 80, 60, 50]$ ▷ facilities opening multiples
  for $t \in T$ **do**
    for $pack \in PACK$ **do**
      for $s \in S_d$ **do**
        for $z \in Z$ **do**
          try to allocate *pack* shots of z on facility s at slots t and $t + \Delta t_z$
          if feasible **then**
            update D_d
            update capacity of s at t and $t + \Delta t_z$
            Break

6.5.1 Use of the MILP model and heuristic algorithm

In order to allow a fair comparison between the MILP model, which was used in practice to schedule the campaigns in Bologna, and the algorithm that mimics a manual solution, the latter was run on the same data and by following the same procedure.

We run each method (MILP or algorithm) starting from the first day of the first campaign, with the real stock availability. For each subsequent campaign which started before the previous one was concluded, the actual availability of vaccine doses and the actual capacity at vaccination centers was computed by deducting the resources allocated to the previous campaign. In this respect, the output of each campaign was used as an input of the following ones. In addition, since some campaigns were performed before those scheduled through the MILP model, and other campaigns were performed in parallel with those considered in this study (e.g., rest homes residents), we removed the corresponding doses from the warehouse availability.

6.6 Results

We implemented the MILP model in the Julia language (Bezanson et al., 2017) with the JuMP package (Dunning et al., 2017), a domain-specific modeling language for mathematical optimization embedded in Julia. JuMP uses a generic solver-independent interface provided by the MathOptInterface package, making it easy to change between a number of open-source and commercial solvers. The resulting model was solved with the GuRoBi MILP solver on a server with an Intel Xeon Gold 6230 processor, 16GB RAM under the CentOS Linux operating system. As already mentioned, our goal is to vaccinate the largest number of people in the shortest time.

During the period between 09/04 and 08/07 a constant evaluation of LHAs efficiency was performed by both the Regional Coordination Center and the National board. Comparing the progress of the LHA of Bologna vaccination campaign with respect to the regional one (Source Open data Commissioner Structure), the vaccinations performed by the LHA covered a share of 20.5%. The number of vaccinations performed by the LHA of Bologna was slightly higher than the supply provided (equal to 19.8% according to the regional distribution criterion of vaccine supplies). The resulting overproduction of 3.5% in the given evaluation horizon can be considered as an efficiency indicator with respect to other regional LHAs. During the same period the National and Regional Coordination Centers provided weekly production targets for each LHA that were, for LHA of Bologna, equal to 606668 vaccinations to be performed. The total number of vaccinations performed by the LHA of Bologna in the given period has been equal to 613568 exceeding the production by 1.1%, also in this case, the vaccination campaign supported by the MILP model proved to be efficient.

Table 6.2 shows detailed information for each considered campaign, comparing the timings obtained with the MILP model and the greedy heuristic algorithm. The table reports the population size, the campaign starting date, the date when the last individual was vaccinated and the mean waiting time (M.W.T). We see that at the beginning the two solutions have a similar mean waiting time but then we have a divergent trend [3].

The performance degradation of the algorithm is due to a domino effect that leads to an accumulation of delay, in particular in presence of second shots and limited vaccine supplies. In the dynamic context of a vaccination

[3] the OVER 60 campaign started before the OVER 65 campaign due to supply and policy changes.

DATA			MILP		ALGORITHM	
Campaign	Population	Start date	End date	M.W.T.	End date	M.W.T.
OVER 70 + VULN.	69126	11/04/2021	15/05/2021	18	21/05/2021	17.8
OVER 60	34962	29/04/2021	27/06/2021	18.4	03/06/2021	18.7
OVER 65	24548	13/05/2021	30/05/2021	7	14/06/2021	15.4
FRAGILE	15752	14/05/2021	06/06/2021	12.3	23/06/2021	20.6
OVER 55	39306	01/06/2021	15/06/2021	4.3	14/07/2021	10.6
OVER 50	30426	08/06/2021	27/06/2021	6.7	22/07/2021	8.8
OVER 40	68756	13/06/2021	20/07/2021	13.7	25/08/2021	42.3
OVER 18	78586	01/07/2021	20/08/2021	23.6	02/11/2021	67.1

TABLE 6.2: Solution comparison for each campaign

campaign, greedy choices, which can look good in the short time, have a negative impact on future opportunities.

Over the whole considered population of 361462 individuals, the average waiting time (in days) for the MILP solution is 60.4 days, while this figure is almost 17 days larger for the algorithm, with an average waiting time of 77.0 days.

Figure 6.3 show how the number of vaccine shots evolves over time. The algorithm solution has higher peaks but it is less regular and has troubles in assigning the first shots for the month of July, mainly for lack of capacity in the vaccination facilities. The MILP solution, instead, is more regular and the whole campain has an earlier termination date.

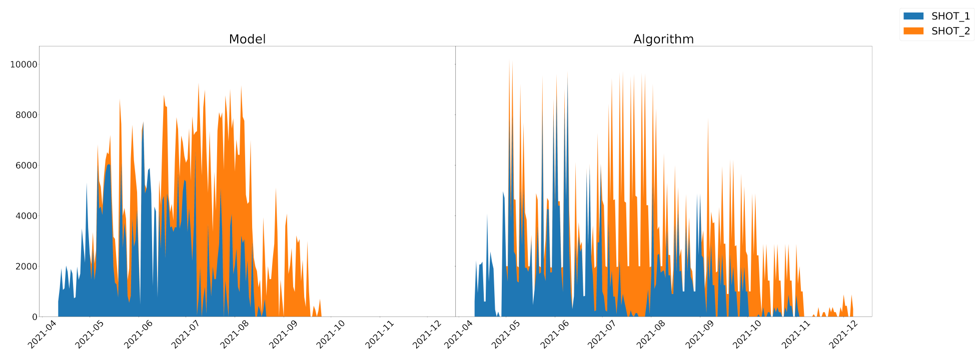


FIGURE 6.3: Vaccine shots trend

Figure 6.4 represents the stock evolution over time for each vaccine from 11/04/2021 to 30/06/2021, and it confirms that the algorithm performance deterioration is not due to stock availability.

The picture show that the consumption rate for Comirnaty is faster in the model solution as soon as there is a delivery. Moreover, Comirnaty trend consumption in the model solution is regular while in the algorithm solution a surplus stock is being accumulated. The stock of Vaxzevria follows a similar trend in both solutions, but in the algorithm solution there is long period of out-of-stock. Finally, Spikevax and Janssen have a limited impact on the vaccination campaign.

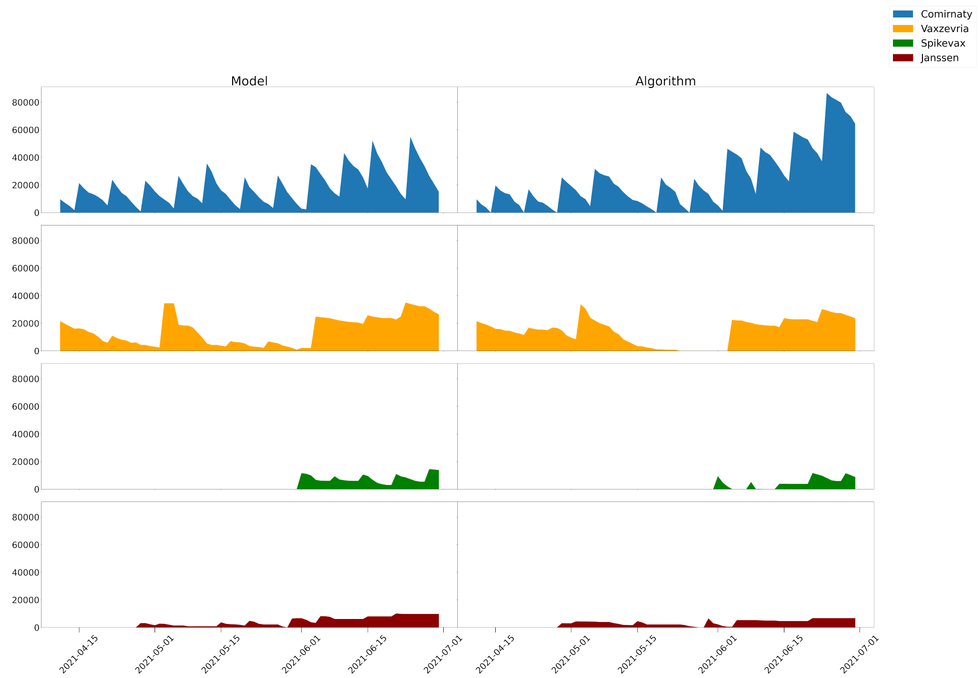


FIGURE 6.4: Stock trend for vaccine

6.7 Conclusions

In this chapter we have described the decisions that need to be taken for organizing a vaccination campaign during a pandemic, subject to strict constraints on availability of vaccine doses and limited capacity at vaccination centers. The design of the campaign is made even more difficult by the mix of different vaccines with diverse second dose delay and target population. In the specific case of the COVID-19 pandemic, the effectiveness of the decision process was jeopardized by the extreme uncertainty concerning not only the forecast of the resources, but also the target population for each vaccine, which in some cases was changed from morning to night.

These difficulties made necessary to develop a fast multi-scenario tool to support the scheduling of campaigns. To this aim, we have introduced an optimization model which gave the decision maker the possibility to test different settings and have a configurable solution within a few seconds. A manual solution would have required about two working days with higher error probability, poor control of all the constraints and low flexibility, and also showed of lower quality in our analysis.

This tool allowed the Bologna LHA to perform over the regional request and face a challenging problem in a emergency condition, minimizing organization times. Currently, this tool is used to support the organization of vaccination of post-second shots. This is certainly a less complex problem, from a computational point of view, but it still represents a priority objective for the LHA.

The introduction of stochastic elements within the model could be a promising direction. The most obvious area for this improvement could be the area of supplies. By introducing such an improvement, it would be possible to make impact forecasts over a longer time span and have more robust solutions, even in different start-up scenarios.

Bibliography

Aboueljinane, L., Sahin, E., & Jemai, Z. (2013). A review on simulation models applied to emergency medical service operations. *Computers & Industrial Engineering*, *66*(4), 734–750.

Agulló, V., Fernández-González, M., de la Tabla, V. O., Gonzalo-Jiménez, N., Garcıóa, J. A., Masiá, M., & Gutiérrez, F. (2021). Evaluation of the rapid antigen test Panbio COVID-19 in saliva and nasal swabs in a population-based point-of-care study. *Journal of Infection*, *82*(5), 186–230.

Alassafi, M. O., Jarrah, M., & Alotaibi, R. (2022). Time series predicting of COVID-19 based on Deep Learning. *Neurocomputing*, *468*, 335–344.

Alkharashi, M. S., Alsharif, H. M., Altahan, F. A., Alrashed, A. W., & Abdulghani, M. (2022). The Impact of the COVID-19 Pandemic on Ophthalmic Outpatient Care in a Tertiary Care Center in Riyadh. *10*(9), 1654.

Araja, D., Berkis, U., & Murovska, M. (2022). COVID-19 pandemic-revealed consistencies and inconsistencies in healthcare: A medical and organizational view. *10*(6), 1018.

Aringhieri, R., Bigharaz, S., Druetto, A., Duma, D., Grosso, A., & Guastalla, A. (2022). The daily swab test collection problem. *Annals of Operations Research*, 1–22.

Aringhieri, R., Bruni, M., Khodaparasti, S., & van Essen, J. T. (2017). Emergency medical services and beyond: Addressing new challenges through a wide literature review. *Computers & Operations Research*, *78*, 349–368.

ArunKumar, K., Kalaga, D. V., Kumar, C. M. S., Chilkoor, G., Kawaji, M., & Brenza, T. M. (2021). Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA). *Applied Soft Computing*, *103*, 107161.

Asplin, B. R., Magid, D. J., Rhodes, K. V., Solberg, L. I., Lurie, N., & Camargo Jr, C. A. (2003). A conceptual model of emergency department crowding. *Annals of Emergency Medicine*, *42*(2), 173–180.

Beaulieu, M.-D., Haggerty, J., Tousignant, P., Barnsley, J., Hogg, W., Geneau, R., Hudon, É., Duplain, R., Denis, J.-L., Bonin, L., et al. (2013). Characteristics of primary care practices associated with high quality of care. *Cmaj*, *185*(12), E590–E596.

Bell, L., & Wagner, R. (2019). Modeling emergency room arrivals using the poisson process. *The College Mathematics Journal*, *50*(5), 343–350.

Benoni, R., Campagna, I., Moretti, F., & Tardivo, S. (2022). Comparing swab- and different symptoms-based strategies to ascertain COVID-19 recovery in healthcare workers: A cost-effectiveness analysis. *Cost Effectiveness and Resource Allocation*, *20*(1), 50.

Bertsimas, D., Digalakis Jr, V., Jacquillat, A., Li, M., & Previero, A. (2021). Where to locate COVID-19 mass vaccination facilities? *Naval Research Logistics*.

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review*, *59*(1), 65–98.

Bhimraj, A., Morgan, R. L., Shumaker, A. H., Lavergne, V., Baden, L., Cheng, V. C.-C., Edwards, K. M., Gandhi, R., Muller, W. J., O'Horo, J. C., et al. (2020). Infectious Diseases Society of America Guidelines on the treatment and management of patients with coronavirus disease 2019 (COVID-19). *Clinical Infectious Diseases*.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the natural language toolkit*. "O'Reilly Media, Inc."

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., et al. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *The Lancet*, *395*(10223), 507–513.

Cho, M., Song, M., Park, J., Yeom, S.-R., Wang, I.-J., & Choi, B.-K. (2020). Process Mining-supported emergency room process performance indicators. *International Journal of Environmental Research and Public Health*, *17*(17), 6290.

Cildoz, M., Ibarra, A., & Mallor, F. (2019). Accumulating priority queues versus pure priority queues for managing patients in emergency departments. *Operations Research for Health Care*, *23*, 100224.

Cildoz, M., Mallor, F., & Ibarra, A. (2018). Analysing the ED patient flow management problem by using accumulating priority queues and simulation-based optimization. *2018 Winter Simulation Conference (WSC)*, 2107–2118.

Cinaglia, P., & Cannataro, M. (2022). Forecasting COVID-19 epidemic trends by combining a neural network with rt estimation. *Entropy*, *24*(7), 929.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding.

Dionisie, V., Ciobanu, A. M., Moisa, E., Manea, M. C., & Puiu, M. G. (2022). The Impact of the COVID-19 Pandemic on Inpatient Admissions for Psychotic and Affective Disorders: The Experience of a Large Psychiatric Teaching Hospital in Romania. *10*(8), 1570.

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*, *20*(5), 533–534.

Drożdżal, S., Rosik, J., Lechowicz, K., Machaj, F., Szostak, B., Przybyciński, J., Lorzadeh, S., Kotfis, K., Ghavami, S., & Łos, M. J. (2021). An update on drugs with therapeutic potential for SARS-CoV-2 (COVID-19) treatment. *Drug Resistance Updates*, *59*, 100794.

Duma, D., & Aringhieri, R. (2020). An ad hoc process mining approach to discover patient paths of an Emergency Department. *Flexible Services and Manufacturing Journal*, *32*.

Duma, D., & Aringhieri, R. (2023). Real-time resource allocation in the Emergency Department: A case study. *Omega*, *117*, 102844.

Dumas, M., La Rosa, M., Mendling, J., & A Reijers, H. (2013). *Fundamentals of business process management*. Springer.

Dunning, I., Huchette, J., & Lubin, M. (2017). JuMP: A modeling language for mathematical optimization. *SIAM Review*, *59*(2), 295–320.

Garnett, L., Bello, A., Tran, K. N., Audet, J., Leung, A., Schiffman, Z., Griffin, B. D., Tailor, N., Kobasa, D., & Strong, J. E. (2020). Comparison analysis of different swabs and transport mediums suitable for SARS-CoV-2 testing following shortages. *Journal of virological methods*, *285*, 113947.

Georgiadis, G., & Georgiadis, M. (2021). Optimal planning of the COVID-19 vaccine supply chain. *Vaccine*, *39*.

Griewing, S., Kalder, M., Lingenfelder, M., Wagner, U., & Gremke, N. (2022). Impact of the COVID-19 Pandemic on Gyne-Oncological Treatment—A Retrospective Single-Center Analysis of a German University Hospital with 30,525 Patients. *10*(12), 2386.

Günal, M., & Pidd, M. (2009). Understanding target-driven action in emergency department performance using simulation. *Emergency Medicine Journal*, *26*(10), 724–727.

Gunasekaran, A., & Kobu, B. (2007). Performance measures and metrics in logistics and supply chain management: A review of recent literature (1995–2004) for research and applications. *International Journal of Production Research*, *45*(12), 2819–2840.

Gupta, D., & Denton, B. (2008). Appointment scheduling in health care: Challenges and opportunities. *IIE Transactions*, *40*(9), 800–819.

Hadj Hassine, I. (2022). Covid-19 vaccines and variants of concern: A review. *Reviews in medical virology*, *32*(4), e2313.

Han, S., Cai, J., Yang, J., Zhang, J., Wu, Q., Zheng, W., Shi, H., Ajelli, M., Zhou, X.-H., & Yu, H. (2021). Time-varying optimization of COVID-19 vaccine prioritization in the context of limited vaccination capacity. *Nature Vommunications*, *12*(1), 4673.

Higginson, I. (2012). Emergency Department crowding. *Emergency Medicine Journal*, *29*(6), 437.

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., & Gu, X. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, *395*(10223), 497–506.

Hulshof, P. J., Kortbeek, N., Boucherie, R. J., Hans, E. W., & Bakker, P. J. (2012). Taxonomic classification of planning decisions in health care: A structured review of the state of the art in OR/MS. *Health Systems*, *1*, 129–175.

Ibrahim, D., Kis, Z., Tak, K., Papathanasiou, M., Kontoravdi, C., Chachuat, B., & Shah, N. (2022). Optimal design and planning of supply chains for viral vectors and RNA vaccines. In *Computer Aided Chemical Engineering* (pp. 1633–1638). Elsevier.

Italian Ministry of Health. (2008). Emergency Department and DEA classification. https://www.salute.gov.it/portale/prontoSoccorso/dettaglioContenutiProntoSoccorso.jsp?lingua=italiano%5C&id=1190&area=118%5C%20Pronto%5C%20Soccorso%5C&menu=vuoto

Jayaprakash, N., O'Sullivan, R., Bey, T., Ahmed, S. S., & Lotfipour, S. (2009). Crowding and delivery of healthcare in emergency departments: The European perspective. *Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health*, *10*(4).

Jun, J. B., Jacobson, S. H., & Swisher, J. R. (1999). Application of discrete-event simulation in health care clinics: A survey. *Journal of the Operational Research Society*, *50*(2), 109–123.

Kaplan, R. S., Norton, D. P., et al. (2000). Having trouble with your strategy? Then map it. *Focusing Your Organization on Strategy—with the Balanced Scorecard*, *49*(5), 167–176.

Król, K., & Zdonek, D. (2020). Analytics maturity models: An overview. *Information*, *11*(3), 142.

Książek, R., Kapłan, R., Gdowska, K., & Łebkowski, P. (2022). Vaccination Schedule under Conditions of Limited Vaccine Production Rate. *Vaccines*, *10*(1), 116.

Kueng, P. (2000). Process performance measurement system: A tool to support process-based organizations. *Total Quality Management*, *11*(1), 67–85.

La Rosa, G., Mancini, P., Ferraro, G. B., Veneri, C., Iaconelli, M., Bonadonna, L., Lucentini, L., & Suffredini, E. (2021). SARS-CoV-2 has been circulating in northern Italy since December 2019: Evidence from environmental monitoring. *Science of the Total Environment*, *750*, 141711.

Lagoe, R. J., & Jastremski, M. S. (1990). Relieving overcrowded emergency departments through ambulance diversion. *Hospital topics*, *68*(3), 23–27.

Lai, C.-C., Shih, T.-P., Ko, W.-C., Tang, H.-J., & Hsueh, P.-R. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and Coronavirus Disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*, *55*(3), 105924.

Leo, G., Lodi, A., Tubertini, P., & Di Martino, M. (2016). Emergency Department Management in Lazio, Italy. *Omega*, *58*, 128–138.

Liu, K., & Lou, Y. (2022). Optimizing COVID-19 vaccination programs during vaccine shortages: A review of mathematical models. *Infectious Disease Modelling*.

Liu, Z., Rexachs, D., Epelde, F., & Luque, E. (2017). An agent-based model for quantitatively analyzing and predicting the complex behavior of emergency departments. *Journal of Computational Science*, *21*, 11–23.

Lu, H., Stratton, C. W., & Tang, Y.-W. (2020). Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle. *Journal of medical virology*, *92*(4), 401.

Mans, R., Van der Aalst, W., & Vanwersch, R. (2015). *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*. Springer International Publishing.

McLean, G., Kamil, J., Lee, B., Moore, P., Schulz, T. F., Muik, A., Sahin, U., Türeci, Ö., & Pather, S. (2022). The impact of evolving SARS-CoV-2 mutations and variants on COVID-19 vaccines. *Mbio*, *13*(2), e02979–21.

Mohamed, K., Rzymski, P., Islam, M. S., Makuku, R., Mushtaq, A., Khan, A., Ivanovska, M., Makka, S. A., Hashem, F., Marquez, L., et al. (2022). COVID-19 vaccinations: The unknowns, challenges, and hopes. *Journal of Medical Virology*, *94*(4), 1336–1349.

Müller, M., Schechter, C. B., Hautz, W. E., Sauter, T. C., Exadaktylos, A. K., Stock, S., & Birrenbach, T. (2021). The development and validation of a resource consumption score of an emergency department consultation. *PLoS One*, *16*(2), e0247244.

Ochani, R., Asad, A., Yasmin, F., Shaikh, S., Khalid, H., Batra, S., Sohail, M. R., Mahmood, S. F., Ochani, R., Arshad, M. H., et al. (2021). COVID-19 pandemic: From origins to outcomes. A comprehensive review of viral pathogenesis, clinical manifestations, diagnostic evaluation, and management. *Infez Med*, *29*(1), 20–36.

Onyeaka, H., Anumudu, C. K., Al-Sharify, Z. T., Egele-Godswill, E., & Mbaegbu, P. (2021). COVID-19 pandemic: A review of the global lockdown and its far-reaching effects. *Science progress*, *104*(2), 00368504211019854.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). Pytorch: An imperative style, high-performance deep learning library.

Paul, S. A., Reddy, M. C., & DeFlitch, C. J. (2010). A systematic review of simulation studies investigating emergency department overcrowding. *Simulation*, *86*(8-9), 559–571.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830.

Petrelli, F., Cangelosi, G., Nittari, G., Pantanetti, P., Debernardi, G., Scuri, S., Sagaro, G. G., Nguyen, C. T. T., & Grappasonni, I. (2021). Chronic Care Model in Italy: A narrative review of the literature. *Primary Health Care Research & Development*, *22*, e32.

Rebuge, Á., & Ferreira, D. R. (2012). Business process analysis in healthcare environments: A methodology based on process mining. *Information Systems*, *37*(2), 99–116.

Rehman, S. U., Rehman, S. U., & Yoo, H. H. (2021). COVID-19 challenges and its therapeutics. *Biomedicine & Pharmacotherapy*, *142*, 112015.

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., Hasell, J., Macdonald, B., Beltekian, D., & Roser, M. (2020). Coronavirus pandemic (COVID-19). *Our world in data*.

Rojas, E., Cifuentes, A., Burattin, A., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2019). Performance analysis of emergency room episodes through Process Mining. *International Journal of Environmental Research and Public Health*, *16*(7), 1274.

Schneider, S. M., Hamilton, G. C., Moyer, P., & Stapczynski, J. S. (1998). Definition of emergency medicine. *Academic emergency medicine*, *5*(4), 348–351.

Sharma, A., Ahmad Farouk, I., & Lal, S. K. (2021). COVID-19: A review on the novel coronavirus disease evolution, transmission, detection, control and prevention. *Viruses*, *13*(2), 202.

Shim, E. (2021). Optimal allocation of the limited COVID-19 vaccine supply in South Korea. *Journal of Clinical Medicine*, *10*(4), 591.

Simeunović, N., Kamenko, I., Bugarski, V., Jovanović, M., & Lalić, B. (2017). Improving workforce scheduling using artificial neural networks model. *Advances in Production Engineering & Management*, *12*(4), 337–352.

SimPy: Discrete event simulation for Python. (2002).

Soria-Arguello, I., Torres, R., Perez, H., & Perea, G. (2021). A Proposal Mathematical Model for the Vaccine COVID-19 Distribution Network: A Case Study in Mexico. *Mathematical Problems in Engineering*, *2021*, 1–11.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tang, L., Li, Y., Bai, D., Liu, T., & Coelho, L. C. (2022). Bi-objective optimization for a multi-period COVID-19 vaccination planning problem. *Omega*, *110*, 102617.

Tavana, M., Govindan, K., Nasr, A. K., Heidary, M. S., & Mina, H. (2021). A mathematical programming approach for equitable COVID-19 vaccine distribution in developing countries. *Annals of Operations Research*, 1–34.

To, K. K.-W., Tsang, O. T.-Y., Yip, C. C.-Y., Chan, K.-H., Wu, T.-C., Chan, J. M.-C., Leung, W.-S., Chik, T. S.-H., Choi, C. Y.-C., Kandamby, D. H., et al. (2020). Consistent detection of 2019 novel coronavirus in saliva. *Clinical Infectious Diseases*, *71*(15), 841–843.

Van der Aalst, W. (2016). *Process mining: Data science in action* (Vol. 2). Springer.

Van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., Tamin, A., Harcourt, J. L., Thornburg, N. J., Gerber, S. I., et al. (2020). Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *New England journal of medicine*, *382*(16), 1564–1567.

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace.

Vance, J., & Sprivulis, P. (2005). Triage nurses validly and reliably estimate emergency department patient complexity. *Emergency Medicine Australasia*, *17*(4), 382–386.

Vom Brocke, J. (2007). Service portfolio measurement: Evaluating financial performance of service-oriented business processes. *International Journal of Web Services Research*, *4*(2), 1–32.

Wang, C., Horby, P. W., Hayden, F. G., & Gao, G. F. (2020). A novel coronavirus outbreak of global health concern. *The Lancet*, *395*(10223), 470–473.

Wang, H., Robinson, R. D., Bunch, K., Huggins, C. A., Watson, K., Jayswal, R. D., White, N. C., Banks, B., & Zenarosa, N. R. (2014). The inaccuracy of determining overcrowding status by using the national ED overcrowding study tool. *The American Journal of Emergency Medicine*, *32*(10), 1230–1236.

Weiss, S., Derlet, R., Arndahl, J., Ernst, A., Richards, J., Fernandez-Frackelton, M., Schwab, R., Stair, T., Vicellio, P., Levy, D., Brautigan, M., Johnson, A., Nick, T., & Fernández-Frankelton, M. (2004). Estimating the Degree of Emergency Department Overcrowding in Academic Medical Centers: Results of the National ED Overcrowding Study (NEDOCS). *Academic Emergency Medicine : Official Journal of the Society for Academic Emergency Medicine*, *11*, 38–50.

Welch, S., & Savitz, L. (2012). Exploring strategies to improve emergency department intake. *The Journal of Emergency Medicine*, *43*(1), 149–158.

WHO. (2022). Ageing and health.

Wu, H.-Y. (2012). Constructing a strategy map for banking institutions with key performance indicators of the balanced scorecard. *Evaluation and Program Planning*, *35*(3), 303–320.

Zhang, C., Li, Y., Cao, J., & Wen, X. (2022). On the mass COVID-19 vaccination scheduling problem. *Computers & Operations Research*, *141*, 105704.