TÉCNICO LISBOA

Behavioural Operational Research approaches for the assessment of surgical voucher acceptance decisions

Madalena Comporta de Almeida

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisors: Prof. Daniel Rebelo dos Santos
Dra. Ana Maria Silvestre Duarte

Examination Committee

Chairperson: Prof. Mónica Duarte Correia de Oliveira
Supervisor: Prof. Daniel Rebelo dos Santos
Member of the Committee: Prof. Miguel Jorge Vieira

November 2021

To my grandmother Isaura,

who every morning took me by the hand to the school bus stop

Declaration

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgments

This dissertation, more than an isolated final essay, is the culmination of all my years at Técnico. As an important stage of my life comes close to an end, it is imperative to thank and appreciate everyone who helped me along the way. After all, I would not have been able to do it alone.

To start, I would like to express my gratitude to Dr. Daniel Santos and to Eng. Mariana Oliveira, who guided me for the past months while I was developing this project. Their weekly advice was crucial, helping me to find clarity in so many moments of doubt and confusion.

I would also like to thank all my friends who, in one way or another, helped me through my university journey. The ones who experienced this adventure of resilience with me, for sharing the same desperate moments and (many more) happy memories – we suffered, but we also had a lot of fun! The ones who were on their separate paths, but kept making the effort to get together during the weekends or the holidays. The ones who, although only in my heart, completed this journey by my side.

To my whole family, thank you for believing in me and being proud of my achievements. I deeply appreciate every advice and love gestures you had with me – Careca, thank you for the carpools; prima Lena, thank you for the sleepovers. A very special thanks to my godparents, who welcomed me into their home for the past years. You were my ”part-time parents” from monday to friday and gave me a sense of home away from home, which I will forever appreciate.

To my parents, I am grateful for every little thing they did – and do – for me. Thank you for the love, support, and all the opportunities you offered me. For every late night you stayed up helping me finish some elementary school’s project I forgot was due the next morning – I think I learned the lesson (at least I completed this dissertation in time).

Resumo

Em Portugal, o acesso a cuidados de saúde tem sido prejudicado por longos tempos de espera para cirurgia. Uma das políticas implementadas para lidar com este problema é a atribuição de vales a pacientes cujo tempo de espera ultrapasse os limites expressos na lei, oferecendo-lhes a oportunidade de serem transferidos para outro hospital. Contudo, este programa de transferências não tem alcançado resultados satisfatórios, pois apenas cerca de um quinto dos vales é aceite pelos pacientes.

Nesta dissertação, foi realizada uma abrangente investigação comportamental para compreender as principais causas das decisões dos pacientes relativamente aos vales – representadas através do *health belief model* (modelo de crenças em saúde), desenvolvido com base em respostas a um questionário, depois de efetuada uma análise de equações estruturais – e estudar o impacto que certas alterações comportamentais podem ter na taxa de aceitação e, consequentemente, nos números da lista de espera –, com o auxílio de um modelo de simulação onde diferentes cenários foram testados.

O modelo de comportamento revelou que as barreiras percecionadas pelos pacientes relativamente ao seu processo de transferência, bem como certos incentivos à ação (de aceitar), foram os fatores determinantes na origem da intenção de aceitar o vale. Para além disso, uma diminuição de 50% ao nível das barreiras percecionadas foi suficiente para produzir uma taxa de aceitação de 75.5% na simulação, verificando-se menos cerca de 32 000 pacientes em espera para cirurgia ao fim de um ano.

Este trabalho contribui com duas ferramentas úteis para as autoridades de saúde, que podem ser usadas para melhorar os resultados da política de vales de cirurgia e os seus efeitos na lista de espera. Mais precisamente, o modelo comportamental pode ajudar na identificação do tipo de medidas com potencial para influenciar o processo de decisão dos pacientes, aumentando a eficácia deste programa; já a simulação, permite testar as consequências práticas de diferentes cenários e intervenções.

Palavras-chave: Lista de espera para cirurgia, Comportamento em saúde, Modelo de crenças em saúde, Simulação de agentes individuais

Abstract

Portuguese citizens' access to healthcare has been compromised by long waiting times for surgery. One of the measures applied to tackle this problem is the attribution of vouchers to patients who have been waiting for a period of time that surpasses the maximum limits established by law, offering them the option to be transferred to a different hospital. However, this transfer programme has not shown satisfactory results, as only about one-fifth of the emitted vouchers end up being accepted by patients.

Within this dissertation, a comprehensive behaviour operational research was conducted, in order to better understand the main causes of patients' decisions regarding these vouchers – represented by the health belief model, that was built based on a questionnaire's answers analysed using structural equation modelling –, and to study the impact that changes of some behavioural key factors could have on the acceptance rate and, ultimately, on the surgery waiting list's numbers – with the help of a simulation model where different behavioural scenarios were tested.

The behavioural model revealed that patients' perceived barriers regarding transfer, as well as cues to action, were the significant determining factors to explain intention to accept the voucher. Moreover, a 50% decrease in perceived barriers was enough to produce a 75.5% voucher acceptance in the simulation, leading to about less 32 000 patients waiting for their treatment after one year.

This work provides health authorities with two valuable tools to improve surgical vouchers' outcomes, boosting their positive effects on the waiting list. Namely, the behaviour model helps aiming public strategies at the psychological factors that determine patients' decision-making, increasing new policies' chances of success; the simulation allows them to test the practical consequences of different scenarios.

Keywords: Surgery waiting list, Health-related behaviour, Health belief model, Agent-based simulation

Contents

List of Tables

List of Figures

Acronyms

ABS Agent-based Simulation

ACSS *Administração Central do Sistema de Saúde* (portuguese for Central Administration of the Health System)

ADSE *Assistência na Doença aos Servidores Civis do Estado* (portuguese for Disease Assistance to Civil Servants)

ARS *Autoridade Regional de Saúde* (portuguese for Regional Health Authority)

AVE Average Variance Extracted

BOR Behavioural Operational Research

CFA Confirmatory Factor Analysis

CFI Comparative Fit Index

CR Composite Reliability

CSH *Cuidados de Saúde Hospitalares* (portuguese for Hospital Care Services)

CtA Cues to Action

CTH *Consulta a Tempo e Horas* (portuguese for Medical Appointments on Time)

DES Discrete-event Simulation

EFA Exploratory Factor Analysis

EURO Association of European Operational Research Societies

FA Factor Analysis

HBM Health Belief Model

HM Health Motivation

ItA Intention to Accept

KMO Kaiser-Meyer-Olkin

LTV Lisbon and Tagus Valley

MHB My Health Book

OECD Organisation for Economic Co-operation and Development

OR Operational Research

PA Parallel Analysis

PBn Perceived Benefits

PBr Perceived Barriers

PECLEC *Programa Especial de Combate às Listas de Espera Cirúrgicas* (portuguese for Special Programme to Combat Surgery Waiting Lists)

PECS Physis, Emotion, Cognition, and Status

PERLE *Programa Específico de Resolução de Listas de Espera* (portuguese for Specific Waiting List Resolution Program)

PMT Protection Motivation Theory

PPA *Programa de Promoção do Acesso* (portuguese for Access Promotion Program)

PPI Patient and Public Involvement

PPP Public-Private Partnership

PTh Perceived Threat

RMSEA Root Mean Square Error Approximation

SAMS *Serviços de Assistência Médico Social* (portuguese for Social Health Care Services)

SCT Social Cognitive Theory

SD System Dynamics

SEM Structural Equation Modelling

SI *Sistema de Informação* (portuguese for Information System)

SIGA *Sistema Integrado de Gestão do Acesso* (portuguese for Integrated System of Healthcare Access)

SIGIC *Sistema Integrado de Gestão de Inscritos para Cirurgia* (portuguese for Integrated System for the Management of the Surgery Waiting List)

SIGLIC *Sistema Informático de Gestão da Lista de Inscritos para Cirurgia* (portuguese for Information System for the Management of the Surgery Waiting List)

SNS *Serviço Nacional de Saúde* (portuguese for National Health Service)

SV Surgical Voucher

SWL Surgery Waiting List

TAM Technology Acceptance Model

TLI Tucker-Lewis Index

TMRG *Tempo Máximo de Resposta Garantida* (portuguese for Guaranteed Maximum Response Time)

TN Transfer Note

TPB Theory of Planned Behaviour

TRG *Tempo de Resposta Garantida* (portuguese for Guaranteed Response Time)

UGA *Unidade de Gestão do Acesso* (portuguese for Access Management Unit)

UK United Kingdom

ULGA *Unidade Local de Gestão do Acesso* (portuguese for Local Access Management Unit)

URGA *Unidade Regional de Gestão do Acesso* (portuguese for Regional Access Management Unit)

VHI Voluntary Health Insurance

Chapter 1
Introduction

Among the biggest concerns of any healthcare system, there must be the provision of timely and responsive services. With citizens' health at the core of concern, efforts should be made to treat everyone according to their necessities, namely regarding response times. In healthcare, long waits can be a catalyst for health deterioration, therefore harming people's quality of life (sometimes irreversibly).

With this in mind, governments have put into action public health policies that attempt to minimize waiting times for a variety of services. Such policies should be subjected to evaluations that assess their efficacy, help identifying possible problems and testing hypothetical adjustments. That is precisely the reasoning behind this dissertation.

This first chapter introduces the problem addressed throughout this project, exploring its underlying motivation (Subchapter 1.1), the main objectives and methodologies employed (Subchapter 1.2), and presenting an overview of this document's structure (Subchapter 1.3).

1.1 Motivation

Excessively long waiting lists and times for surgery have been a problem of the Portuguese healthcare system over the years, compromising one of citizens' most fundamental rights – the access to timely care treatments. Since 2004, the management of the national surgery waiting list has been executed based on an integrated system, SIGIC, where reasonable waiting time limits for different levels of patients' clinical priority are defined. When a certain hospital does not have enough delivery capacity to schedule a surgery within those limits, the law dictates that a transfer offer must be issued to the patient of concern – either in the form of a surgical voucher or of a transfer note –, so that he/she can be treated at another facility. The latter is, in fact, a particular case of the former (while a surgical voucher can include both public and private providers, a transfer note only offers patients the option to be transferred to other public hospitals). Data from the Ministry of Health usually does not differentiate between these two. Thus, throughout this dissertation, both terms will be often used interchangeably, or, for a matter of simplicity, surgical voucher will most times be used as a reference to both transfer methodologies.

Despite the potential that the surgical vouchers' policy could have had at reducing patients' waiting times for surgery, by transferring part of the high demand from public to private healthcare services, it has fallen short to the expectations. The main reason seems to be a very low acceptance rate among the general public, since the big majority of patients either refuses or ignores this transfer option, preferring to wait longer for treatment at their home hospital. Therefore, there is an urge to explore and understand patients' behaviour and decision-making towards the surgical vouchers' system, so that health authorities may be enlightened regarding possible impacting reforms, capable of increasing vouchers' acceptance level and its effects on the national waiting list.

1.2 Objectives and Methodology

The fundamental aim of this dissertation is to comprehensively explore the surgical vouchers' system implemented in Portugal, more specifically the low number of acceptances registered over the years – from the factors that determine patients' decisions regarding a possible transfer, to the impact that certain behavioural changes can have on this system's performance. To do so, two operational research techniques will be applied in this research: modelling and simulation.

Modelling enables the representation of individual health-related behaviour and the identification of its main determining factors. As an initial step, a psychological model (Health Belief Model) will be adapted to the matter of study. Research hypotheses regarding the nature of the causal associations present in that model will then be stated and tested, based on the results of a proper measurement instrument that will also be developed in this dissertation. Such instrument is intended to measure patients' predisposition to accept a surgical voucher, as well as its determining constructs. The predefined hypotheses validated by these answers will be kept in the final causal model developed to explain intention to behave. The results will also reveal the weights of that causal relationships, showing which are the key factors that truly influence behaviour and that, consequently, should be targeted by health authorities in order to induce behavioural changes among patients.

Simulation is useful to observe the impact that certain changes in the identified causes of behaviour can have on the overall waiting list's statistics. By mimicking the waiting list's environment and patients' possible pathways during their waiting process, it will be possible to analyse multiple virtual scenarios and evaluate how they would influence surgery delivery indicators, such as SVs' acceptance level or the number of surgeries performed annually.

1.3 Outline

This document starts with a contextualization of the problem at hand, in Chapter 2. Chapter 3 provides a review of the available literature regarding behavioural OR applied to healthcare and explores some of the studies where the methodology here employed was based on. Such methodology is explained, in detail, in Chapter 4. It is over at that chapter that the quantitative study of patients' a priori intention to accept a surgical voucher is designed, as well as its corresponding base model, and the final simulation model. Chapter 5 presents the obtained results and discusses their meaning and validity. Finally, in Chapter 6, the conclusions drawn from this research are exposed, as well as some of the limitations faced during its development. Some suggestions for future studies are also mentioned in this final part of the dissertation.

Chapter 2
Context

This chapter aims at providing some context regarding the surgical vouchers' lack of acceptance problem. To do so, it starts with an overview of the Portuguese healthcare system's functioning and organisation (Subchapter 2.1), followed by a description of surgery waiting list's management strategies (Subchapter 2.2), and, lastly, an evaluation of SIGIC's impact on waiting times (Subchapter 2.3).

2.1 Portuguese healthcare system

The Portuguese healthcare system is formed by three co-existing systems: the national health service (SNS, as in *Serviço Nacional de Saúde*), health subsystems (group insurance schemes), and private voluntary health insurance (VHI) [1]. Although there is some overlapping regarding a few of the care services provided by the three of them, the healthcare system is strongly based on SNS, with the other two structures having essentially a supplementary role to the care provided by the former, rather than representing an entirely alternative to it [1].

Thus, the Portuguese health system follows the Beveridge model initially implemented in the UK, where the government acts as a social insurer, assuring healthcare access through a network of public facilities and care providers that form the SNS. Private institutions also play a significant role, although more focused on providing diagnostic, therapeutic and dental services, performing laboratory tests, medical imaging and selling pharmaceutical products. Ambulatory appointments, rehabilitation and hospitalization are also included in the private providers' scope [2]. Access to some of these private care services, namely elective hospital treatment and ambulatory visits, is facilitated by private insurance schemes which, therefore, extend the variety of services offered to citizens already covered by the public healthcare system [2]. Recent health policies have tried to tighten the relationship between public and private providers in order to promote a more efficient healthcare delivery system, which would ultimately benefit the overall health status of the population.

2.1.1 SNS

The SNS is a universal tax-financed healthcare delivery system, guaranteed by the Portuguese state to all residents in its territory. It was initially created in 1979, five years after the revolution that liberated the country from a 41-year-old dictatorship, establishing a democratic regime. This new political framework recognized the right to health as fundamental to everyone, independently of their socioeconomic status, and declared the protection of all citizens' health as responsibility of the state government – which, therefore, should be the main care provider, instead of private institutions, such as religious charities, or the social welfare system previously in place [2].

The SNS became a unifying solution of a set of formerly fragmented and uncoordinated subsystems, mostly inefficient to accomplish health policies' objectives and to achieve significant public health improvements [2]. That was reflected in the poor indicators Portugal had back then – for example, in 1960, child mortality rate reached 77.5%, while 30 years later the registered value was only 10.9% [3].

Although the SNS had been initially declared as "free-of-charge", a constitutional amendment in 1989 changed that statement to "tendentially free", as some user charges were then introduced. Nevertheless, exemptions were also granted in order to protect the initially stated universal right to healthcare and the SNS is still considered almost free at point of delivery [2]. In fact, most of its financing is made through general taxation (either on income or other indirect taxes) and a minor part is guaranteed by small fees (when compared to the total cost of care services) charged to the citizens, following a cost-sharing approach. These user charges usually assume the form of co-payments, where a fixed amount is charged for a certain service, or co-insurances, generally related to drug prescription, where patients pay a fraction of the cost. As stated above, certain patient groups (e.g., citizens with insufficient means, unemployed, people who suffer from chronic diseases, refugees, minors) are exempted from at least some of these out-of-pocket payments in order to protect their access to healthcare. Data from 2016 showed that approximately 60% of the Portuguese population was exempted from user charges [1].

2.1.2 Health subsystems and VHI

Today's health subsystems are remnants of the social welfare system previously installed in the country, according to which access to healthcare was provided based on professional memberships and financed through contributions from both employers and employees. They were not extinguished after the establishment of the SNS, since many trade unions defended them as a way of further protecting their members' access to healthcare, by offering a wider choice of providers than the one presented by the SNS [2]. There are health subsystems established for both public and private workers and their families: ADSE, for civil servants, and SAMS, for banking employees, for example.

Private insurance has been increasing in Portugal, over the years. Nevertheless, it is still a small market, especially when compared to other European countries whose healthcare systems follow a Bismarck model. VHI has its own problems – such as clients' discrimination (attempting to exclude high-risk patients) and the implementation of some exclusions and limitations in the services covered to prevent moral hazard (an increase of healthcare use by insured patients as a result of low out-of-pocket payments) – which suggest that the SNS is still vital to assure universal access to healthcare [1].

Approximately 25% of the population is covered either by a health subsystem or VHI (or both), proving the importance of these supplementary insurance plans to the functioning of the health system as a whole. It is possible for the same person to benefit from triple (or higher) coverage, provided by the

SNS, a health subsystem (that can be doubled if there is a spouse who also benefits from another subsystem) and a VHI [2]. Nevertheless, as these co-existing coverages many times offer overlapping services, it is not very common for a person to have more than a double coverage – adding a health subsystem or a VHI to the SNS services.

2.1.3 Administrative and management structure

Health policies' development and supervision is a task of the central government, more specifically, the Ministry of Health. This office is also responsible for the regulation, planning and management of the SNS and for the regulation, auditing and inspection of private providers. This ministry is subdivided in different management structures with distinct responsibilities. Concerning the budget of SNS, it is defined by the Ministry of Finance and is then allocated to distinct areas by the Ministry of Health [1].

Regarding the SNS, there are two important institutions that act under the government's indirect administration: the ACSS (*Administração Central do Sistema de Saúde*) and five ARSs (*Administrações Regionais de Saúde*). The former is a central organization responsible for managing financial, human and material resources of the SNS. ARSs are distributed along the country (North, Centre, Lisbon and Tagus Valley (LTV), Alentejo, and Algarve) and are in charge of the strategic management of the health of their regions' population, as well as of supervising and controlling their respective hospitals and primary care centres. In addition, both the ACSS and each ARS are responsible for the implementation, regulation and planning of the centrally-developed health policies. For instance, ARSs are in charge of establishing agreements with private providers, so that SNS patients can be treated at those facilities in case the SNS does not possess sufficient means to offer them adequate care [1].

2.2 Problems in healthcare access – the case of elective surgeries

The SNS was created by the Portuguese government to guarantee universal access to healthcare services. This fundamental right designates citizens' possibility to obtain appropriate (i.e., timely and effective) and valuable care services, according to their needs [4]. However, mainly due to inefficiencies detected in the healthcare system, access is sometimes dangerously compromised, which might result in a serious deterioration of the general population's health levels.

Regarding elective (non-urgent) hospital care, two frequently used indicators to evaluate access are: the number of patients on the waiting list for a speciality medical appointment and the number of patients on the waiting list for surgery [5]. Perhaps more important than the length of both lists, is the average time each patient has to wait for the system to answer his/her medical needs and, therefore, waiting times are also vital to take into account when assessing hospital healthcare access. In fact, waiting time is the main factor that influences health outcomes – if patients have to wait a significant amount of time to be treated, their health condition will likely get worse in the meantime (diminishing quality of life and affecting citizens' ability to work, for example). This delay between the moment when the need for surgical treatment is detected and the moment the procedure is executed is not only a burden for the patient but also for the healthcare system itself, since during waiting time usually occurs an overconsumption of SNS services to monitor and control the situation at hand – more medical exams, appointments and drug prescriptions [4].

On the other hand, the fact that there is a considerably long list of patients waiting for a certain procedure, per se, does not imply long waiting times, since a highly efficient and resourceful health system could be able to rapidly answer a big volume of demand. Thus, it seems that the core of the waiting lists' problem concerns the system's efficiency (or lack thereof) in providing care services, rather than the size of those same lists [6]. Nevertheless, in a system with stagnant capacity, longer lists will consequently create higher waiting times.

Waiting lists emerge from a discrepancy between demand and supply: the demand for healthcare services is much superior to the capacity of health providers to attend those needs. Regarding surgery, demand is defined by the number of patients on the waiting list for this procedure, while supply is determined by available capacity, in both the SNS or private providers with which the SNS has previously established agreements to perform those surgeries [4].

In Portugal, a patient is enrolled on the waiting list after a speciality medical appointment where it is detected the need for surgical treatment and after the consent from that same patient. Leaving the list implies that the surgery was performed, either at the SNS or at a private hospital, or the problem was meanwhile solved through another approach. Other reasons might be related to withdrawal or even the death of the patient [4, 5].

Increasing longevity is one of the reasons why healthcare demand is constantly growing – we are living longer and, therefore, experiencing more health problems throughout our lives. At the same time, medical and technological evolution over the years allows for the resolution of many of those problems, often implying surgical interventions. In addition, the security of this procedure has been improving, making it a very common treatment choice nowadays [6, 7].

Delays in healthcare access are a global problem, identified in many other countries besides Portugal. A study revealed that poor efficiency regarding waiting lists' management is more frequent in countries that have a public health insurance system (similar to the SNS) and, simultaneously, that show low surgery response capacity [6]. Healthcare systems analogous to the Portuguese one seem more susceptible to this problem, mainly because of the low out-of-pocket payments, which may increase demand, and due to a lack of delivery capacity, often affected by poor investment and budget cuts.

2.2.1 Waiting lists' management strategies

For public healthcare providers, it may be beneficial to maintain small waiting lists for some of their non-urgent services. Capacity limitations arise as a possible solution to avoid healthcare request of indefinitely increasing, especially in a system where out-of-pocket payments are low and do not help restricting demand [6, 7]. Besides, if a hospital sets its capacity to maximum levels, chances are that many of those resources would stay unused in case demand lowers to moderate levels, resulting in idle time costs. Moreover, and even if demand does not decrease, such investment in delivery capacity often implies unbearable expenses [7].

By maintaining a small number of patients waiting for a given elective medical procedure, providers should be able to guarantee reasonable waiting times (not harmful for patients' health condition) and, at the same time, help hospitals' cost containment [6, 7]. It is important to mention that this approach is only feasible until the point where lists' dimension do not compromise health outcomes as a consequence of excessively increased waiting times. However, maintaining this equilibrium is typically difficult. Consequently, governments all over the world have tried to implement a variety of health policies aimed at minimizing waiting lists and times, often without satisfactory outcomes.

Given that long lists are a result of demand largely surpassing supply, measures could be applied to act on both sides of the problem to obtain optimum waiting times and surgery rates. If the volume of surgeries performed is considered adequate, efforts should be made to contain demand, by setting a priority scale and ordering patients on the list according to it, or by promoting the adoption of VHI, which would transfer part of the demand to private providers. On the other hand, in cases of a low volume of performed surgeries, capacity should be increased, either directly, by investing in more staff (surgeons, nurses) and resources (beds, equipment), or indirectly, through the endorsement of extra activity in public hospitals (following a pay-for-service approach) and the establishment of agreements with the private sector to provide part of the required care services [6].

Evidence has shown that, on a first attempt, governments tend to focus on the supply-side of the issue, applying measures to directly increase the overall capacity to deliver surgical treatments. Many countries started tackling the waiting lists' problem by assigning monetary incentives to extra surgical activity, in order to rapidly increase productivity, over a specific period of time. Such type of policy usually caused a temporary increase in activity (higher number of surgeries performed), but only modest results were obtained regarding waiting times – these either stagnated or slightly decreased. After this initial effect, the improvements obtained with these policies usually fade away. The cause seems to be a positive response of demand to increases in supply – although more people were leaving the waiting list, the number of patients continuously being added was still growing –, which suppressed any positive outcomes. Furthermore, strategies centred on overtime productivity have another disadvantage, as hospitals might feel encouraged to shift ordinary activity to extra time in order to receive monetary rewards. Ireland, for example, did not obtain significant results with a similar incentive programme. However, countries that combined incentives for extra activity with rewards for waiting time reduction, like Spain, were able to achieve their goals. On the other hand, Denmark expanded public services to directly increase supply capacity, also with positive outcomes [6].

However, lasting reductions in waiting times tend to be associated to a better management of the demand-side of the problem, by promoting a more efficient use of existent capacity [6]. To do so, many countries (e.g., Finland, United Kingdom, Denmark, Portugal) established maximum times defined as acceptable for patients to wait for a certain heath service (a surgery, or an appointment, for example) and implemented some kind of priority system, organizing patients according to their clinical condition, instead of on a "first-come-first-served" basis. Nevertheless, the approach used when applying these time limits is not the same across different countries: while some countries established penalties for providers who fail to meet such deadlines, others used them as a strategy to increase patients' central role (they have the possibility to choose providers which guarantee faster treatment) [8].

More often than not, governments put in place a mix of these strategies, aiming at reducing waiting times from both sides of the problem (demand and supply). OECD reports [6, 8] have been analysing the results obtained in multiple countries, from where important conclusions can be drawn – perhaps the most important of all, that temporary measures only lead to temporary results. It is also important to remember that the efficiency of implemented programmes largely depends on the type of healthcare system established in a given country [8].

2.2.2 Portuguese reforms

From a historical perspective, Portugal has had difficulties controlling healthcare access indicators, such as surgery waiting lists. The SNS often struggles to provide an adequate and timely response to the population's problems. Over a ten-year period, there were four different programmes implemented by the government, in successive attempts to decrease the size of surgery waiting lists – PERLE (*Programa Específico de Resolução de Listas de Espera*) in 1995, PPA (*Programa de Promoção do Acesso*) in 1999, PECLEC (*Programa Especial de Combate às Listas de Espera Cirúrgicas*) in 2001, and more recently, SIGIC (*Sistema Integrado de Gestão de Inscritos para Cirurgia*) in 2004 [4].

PERLE, PPA and PECLEC focused essentially on increasing hospitals' capacity of performing surgeries by offering economic stimulus to extra activity (pay-for-service approach), carried out after the normal planned schedule. Besides that, some of these programmes also included measures to encourage collaboration between public and private services, in cases where the SNS itself was not proving to be sufficient to meet demand. However, the majority of these measures failed to achieve their expected objectives – either because of financial difficulties, lack of resources or poor strategical implementation. A major drawback was also the fact that policymakers tried to solve the problem at that specific moment, applying temporary actions and neglecting the long-term nature of the issue. For instance, there was no integrated information system that efficiently monitored waiting lists' evolution over time and the available data was often out of date and unreliable.

Despite the efforts, excessively long surgery waiting lists remained a problem. For instance, PECLEC was able to increase surgical activity at a national level but that increase did not ultimately result in a reduction of waiting times. A possible cause for this lack of improvements in the overall scenario was the diversion of hospital resources from ordinary to overtime activity, which, according to the Portuguese Audit Office, occurred in some hospitals, as a result of the monetary incentives applied at the time [9].

In 2004, a different strategy was put in place. More than a temporary programme to accomplish short-term improvements in waiting lists' length, SIGIC aimed to be an integrated information system, as its name suggests, to allow a better monitoring and management of the whole universe of elective surgeries performed at the SNS, based on updated and reliable data. This programme enabled a better distribution of surgical demand by care providers, taking advantage of the already installed capacity. Thus, patients enrolled on overcrowded waiting lists were offered a transfer opportunity to a different facility (public or private) which presented a faster response. Therefore, through a more efficient management of enrolled patients, waiting times could be reduced and SNS patients' access to care would remain secure [4, 10].

In recent years, efforts have been made to continuously monitor patients' access to all levels of care. SIGA (*Sistema Integrado de Gestão do Acesso*) was created in 2017 and it is responsible for collecting and connecting information otherwise disperse, from other independent information systems, such as SIGIC and CTH (*Consulta a Tempo e Horas*) – the former linked to surgery (as already explained) and the latter to first speciality appointments. These were incorporated in two subsystems of SIGA: SIGA CSH (*Cuidados de Saúde Hospitalares*), related to hospital care access and where SIGIC is included, and SIGA *Primeira Consulta*, which regulates access to speciality care. Besides these, SIGA includes more offices, associated with other types of healthcare services. Thus, access to essential services and referencing between different facilities and levels of care are horizontally monitored and regulated, following patients' path along all parts of the healthcare system. SIGA translates a more patient-centred management strategy, guaranteeing a timely, effective and integrated response of all services [11]. Nevertheless, regarding elective surgery, SIGA CSH follows the previously implemented SIGIC's regulation, which will, therefore, be explored next.

SIGIC's legislative evolution

Created in 2004 and approved by Council of Ministers Resolution no. 79/2004 [10], of June 3^{rd}, this system was firstly implemented in ARSs of Alentejo and Algarve, as a test, and then extended to the rest of the country in the following year [4]. Its main goal was to decrease the time interval between a patient being referred to undergo surgery and the actual moment when that surgery was performed, moving towards a scenario where all surgical treatments take place within clinically acceptable deadlines. SIGIC also intended to be universal, comprising the whole surgical activity happening in the SNS.

SIGIC is associated to an informatic system, SIGLIC (*Sistema Informático de Gestão da Lista de Inscritos para Cirurgia*), which integrates data extracted from hospitals' own information systems.

According to SIGIC's guidelines, the management of a patient who needs surgical care includes the following steps:

- After a referral from a primary care physician, the patient is examined by a speciality doctor at the hospital. There, surgery may be suggested as the most suitable treatment for that patient. If he/she agrees on the proposed procedure, consent must be formalized. The patient is then inserted in the corresponding hospital's waiting list and is given an enrolment certificate, as he/she waits to be called in for surgery, within a time limit previously established;
- In case the hospital does not have enough capacity to perform surgery before the defined time limit, two scenarios can occur: the emission of a transfer note (TN), according to which the patient should be transferred to a SNS hospital nearby that would be able to timely perform the needed procedure; if that is also not possible, the patient is offered a surgical voucher (SV) to use in a public or private hospital of his/her choice, from a list of options (with which the SNS had established agreements), where the intervention will then be executed.

These specific time limits for hospitals to provide a certain service to patients are called TMRGs (*Tempo Máximo de Resposta Garantida*) and, regarding SIGIC's scope, constitute the maximum number of days considered acceptable for patients to wait for a surgical intervention. These limits should be respected by all SNS hospitals, as well as private institutions which had signed contracts with the public healthcare system, in order to assure equality of rights between patients independently of where they are being treated. Nevertheless, hospitals are free to set their own time limits for care delivery, TRGs (*Tempo de Resposta Garantida*), as long as those are lower than the maximum targets (TMRGs) imposed centrally, by the Ministry of Health. Patients also have the right to access information regarding TMRGs as well as to complain to the Health Regulatory Agency in case these are not met [4].

Ministerial Order no. 1529/2008 [12] defined TMRGs according to four different levels of clinical priority (assigned by the speciality doctor who proposes surgery), which allowed patients' differentiation according to their health condition – type of pathology, severity, progression, impact on life quality and expectancy, among other factors. Since 2008, these values have been updated and their most recent version was established by Ministerial Order no. 153/2017 [13], which defines, regarding elective surgery: clinical priority of level four, which refers to deferred emergencies, has a TMRG of 72 hours; level three priority patients have a corresponding TMRG of 15 days; level two has a TMRG of 60 days (that period shortens to 45 days for oncologic and heart disease patients); and, finally, a TMRG of 180 days (60 days for oncologic patients and 60 days for who suffers from a cardiac condition) is defined for priority level one. Hospitals must treat all patients within these TMRGs. When it is not possible, a TN/SV mechanism is put in place to avoid TMRG's surpassing. In those cases, patients should be transferred to other hospitals, according to time limits also established by law:

– For level one priority patients, if after 75% of TMRG (135 days of waiting, for general pathologies) the surgery has still not been scheduled;
– For level two priority patients, if after 50% of TMRG (30 days of waiting, for general pathologies) the surgery has still not been scheduled;
– For priority level three, at patients' request if 25% of TMRG (5 days of waiting) have passed and the surgery has still not been scheduled;
– Patients with clinical priority of level four must be treated as soon as possible and are, therefore, considered untransferable.

Again, these deadlines may be shortened for patients with oncologic or cardiac diseases. This information is summarized below, in Table 1.

Amendments made in 2008 by Ministerial Order no. 45/2008 [14] to SIGIC's regulation conceded patients the right to refuse their transfer to another hospital. The consequences of such decision will be explored further ahead.

SVs give patients the opportunity to choose a public or private facility where they want to be treated, from a list of options. Depending on the kind of contract established with the SNS, private hospitals can be classified into two categories: *hospitais convencionados*, which receive SNS patients previously enrolled on a public hospital's waiting list and who were close to reach their corresponding TMRG, and

Table 1: TMRGs and maximum waiting times until surgery scheduling, according to different levels of clinical priority and types of pathologies [4, 13].

Clinical Priority Level	Type of Pathology	TMRG	Maximum waiting time until scheduling
1	General	180 days	135 days
1	Cardiac	90 days	68 days
1	Oncological	60 days	45 days
2	General	60 days	30 days
2	Oncological/Cardiac	45 days	23 days
3	General/Oncological/Cardiac	15 days	5 days
4	General/Oncological	72 hours	As soon as possible

hospitais protocolados, that treat patients directly forwarded to them from SNS primary care services and who are directly enrolled on that hospital's surgery waiting list [5].

With the incorporation of SIGIC in SIGA, SIGLIC's information was also joined in a larger system, SIGA SI. Although SIGA induced some organizational changes, SIGIC's directives remained in place, as part of SIGA CSH, as previously mentioned [15].

SIGIC's management

SIGIC is a direct responsibility of ACSS, the accountable entity for managing and monitoring patients' access to healthcare. Other institutions within the Ministry of Health, such as ARSs, also play an important role in the system's well-functioning, for instance in the establishment and coordination of agreements with private providers which treat SNS patients across the different regions of Portugal [4].

SIGIC's management organization, as a part of SIGA, is divided into three levels: central, regional and hospital. UGA (*Unidade de Gestão do Acesso*) is the specific unit of ACSS accountable for the implementation and monitoring of all processes essential for SIGIC's efficiency at a national level. It is divided into smaller regional units, URGAs (*Unidade Regional de Gestão do Acesso*), belonging to each ARS. URGAs are responsible for monitoring and controlling the whole process at a regional level, ensuring not only the well-functioning of the system but also that surgical activity meets the expected capacity in hospitals of their geographical area. More specifically, each hospital also has a ULGA (*Unidade Local de Gestão do Acesso*), which role is to guarantee the correct transfer of that hospital's surgical activity information (including patients on the waiting list) to the central system [4]. All three of these levels must cooperate in order to guarantee a correct implementation of SIGIC. For example, regarding patients' transfer process: ULGA is responsible for identifying patients eligible for transfer, as well as to clarify them regarding the whole process; URGA monitors and controls transfers between different healthcare institutions, ensuring services' quality and stakeholders' compliance with the defined protocols, while also ensuring patients' access to information; centrally, UGA selects the list of patients to be transferred and issues their corresponding TN/SVs [4].

Along the process, these units also have the responsibility to act on every time SIGIC's guidelines are disrespected. For instance, when a hospital with which the SNS has previously established a contract to treat transferred SNS patients refuses to comply with that agreement without a plausible justification, ULGA has the authority to suspend the connection with that misconducting hospital [4].

Patients' pathway in the Portuguese surgery waiting list according to SIGIC's regulation

SNS patients' access to hospital speciality appointments is usually dependent on a referral from a general physician at primary care services. Nonetheless, the indication for a speciality appointment can also come from another SNS hospital, private and social institutions, or even as a result of patients' own initiative. Access to this type of secondary care is also regulated by law and restricted by fixed TMRGs – acceptable time limits to wait for the scheduling of first speciality medical visits, defined for three different levels of clinical priority [13].

Once at the designated hospital, a series of medical events centred on a specific health problem is initiated, which is named episode of care. According to the Medical Dictionary for the Health Professions and Nursing [16], an episode of care is defined as "all services provided to a patient with a medical problem within a specific period of time across a continuum of care in an integrated health care system". SIGIC includes all episodes of care that contain surgery.

Patients' consent is essential for the process to progress and, only then, a certificate is issued to formalize patients' enrolment on that hospital's surgery waiting list. This establishes the beginning of a complex process, comprising not only surgery itself but also pre- and post-operative periods, and during which a patient's episode of care must be correctly handled by the healthcare system. In order to assure all occur as smoothly and efficiently as possible, SIGIC's regulation identifies all the different stages and steps of a surgical episode of care (from the first speciality appointment until that episode's conclusion) and provides guidelines explaining what measures should be taken at each step of the way [4].

After the initial referral to speciality care, the following stages of the process essentially consist of: proposal, execution, follow-up and conclusion of the episode at hand. Given the context of this project, the following analysis will be focused on episodes of care related to surgical interventions – where SIGIC's regulation is implemented. Figure 1 represents the above-mentioned sequential steps, as well as their correspondent clinical events.

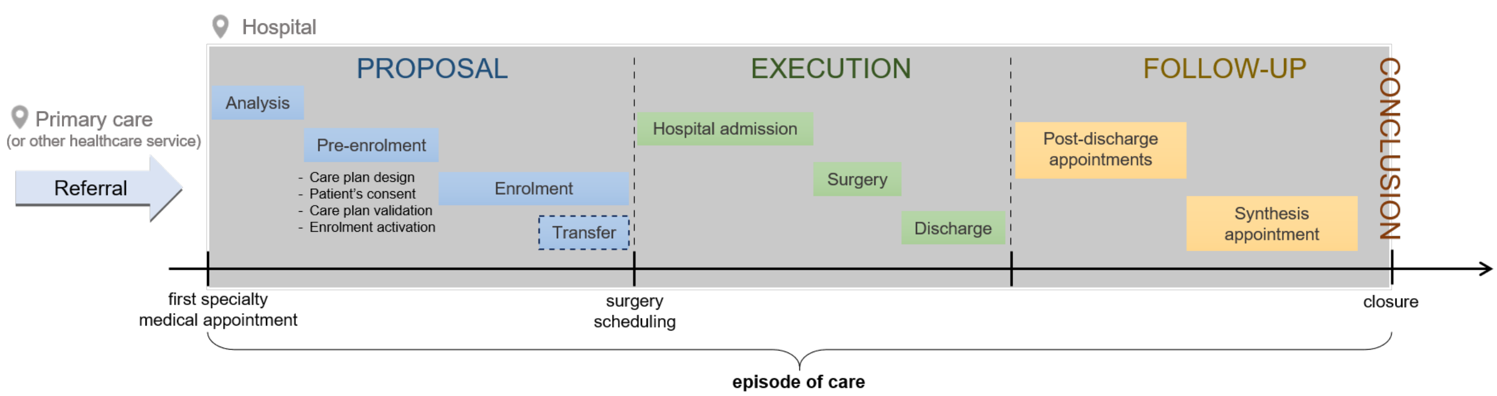

Figure 1: Major stages of a surgical episode of care [4].

The overall process is far from being simple, as many different scenarios can occur during each of the identified stages. Some possible pathways will be explored in the next sub-chapters [4].

Proposal

The proposal is the initial phase of patients' path in hospital care, starting with their first speciality medical appointment. It comprehends all the clinical steps from that moment on and until before the execution of the first key event scheduled for patients' treatment (in this case, surgery). A certain episode of care might actually include more than one key event (multiple surgeries, for instance), but they should all be correlated to the same health problem.

Analysis is the first sub-step of the proposal stage and includes a variable number of events where an extensive evaluation of patients' condition is carried out, in order to collect important data to support a future therapeutic strategy. As proposal's second step, pre-enrolment includes the development of a care plan (which contains the surgical intervention), but also the process of obtaining patients' consent as well as the endorsement of that specific surgical unit's manager. During the enrolment, patients' names enter the waiting list – the registry should be made in the hospital's information system by the corresponding speciality doctor who proposed the intervention and it must specify patients' clinical priority level. Waiting time starts counting the moment patients' care plan is inserted into the system. If all previous steps occur without any problems, an enrolment certificate is emitted by SIGLIC and sent to the patients, confirming their enrolment activation on the waiting list.

As the process advances, patients' priority level can (and should) be adjusted if needed. At any time, the process can be interrupted – for example, cancellation may happen if further medical examination reveals there is, in fact, no need for a patient to undergo surgery; additionally, due to hospital's lack of resources or delay to provide some specific treatment, the process may be transferred to another healthcare facility.

While enrolled on the list, patients may have different features assigned to their process: suspended – an exceptional mechanism during which the process is stagnated and cannot evolve (it can neither be scheduled nor transferred, for example), usually due to technical or logistic issues of the hospital's responsibility –, pending – if patients are not available to undergo surgery at that moment, due to either personal or medical reasons –, untransferable – when the hospital considers a possible transfer would be harmful to a patient, based on clinical or social motives –, or scheduled. Except for suspension and pendency, all the other characteristics can coexist at a certain time for the same registered patient.

Surgery scheduling consists of assigning a date, a time and a specific operating room for a given set of surgeries extracted from the hospital's waiting list. Hospital surgical departments' managers have the responsibility to select a list of patients, based on SIGLIC, to schedule for a specific subsequent time horizon (usually, a day or a week, depending on the planning strategy adopted by the hospital). Patients' selection must obey the following criteria: firstly, patients' clinical priority level and, secondly, antiquity on the list. In case of equal priority levels, patients are ordered according to their antiquity on the waiting list and their respective TMRG (for example, priority level one patients might have different TMRGs according to their type of disease) – favouring those who have been waiting longer and are closer to reaching 100% of TMRG. Once the list of patients to schedule in a given time horizon is defined, that information should be registered in the hospital's informatic system and patients' notification process can be initiated within the deadlines determined by law.

The TMRGs displayed in Table 1 should always be respected, which many times require a transfer process. If a patient's waiting time already surpassed the TMRG, the hospital cannot schedule surgery more than 20 days in advance, so that it does not become an obstacle to a possible transfer.

Transfer

Patients' transfer can be caused by different reasons. The main one is the transferring hospital's lack of delivery capacity to provide surgical care within the TMRGs expressed by law. In addition, specific types of surgery may actually be absent from the set of procedures performed there, implying patients' transfer to other facilities. Independently of the circumstances, transfer only occurs with patients' agreement.

Transfer between different hospitals can occur by mutual agreement, when the hospital where a patient is initially followed concludes that it is not suitable to proceed with the treatment and directly refers the process to another institution, or within the scope of especial programmes, as it happens with the emission of TN/SVs.

According to the situation at hand, different types of transfer can also be carried out. For instance, transfer may respect only the surgical procedure, or it may refer to the whole episode of care, which becomes then an entire responsibility of the receiving hospital. In the first case, post-operative treatments and follow-up remain a responsibility of the hospital from where the patient was transferred. Exceptions include complications identified up to 60 days after hospital discharge, which should then be resolved in the hospital where the surgery was performed. When the whole episode of care is transferred, the receiving hospital becomes, then, in charge not only of surgery itself, but also of the subsequent post-operative treatment necessities. Besides patient's consent, transfer also needs to be approved by the receiving hospital's department manager. Once it is obtained, the TN/SV is considered accepted and the transfer process can be carried out.

As already explained, patients' transfer to other hospital facilities, belonging or not to the SNS, through the emission of TNs or SVs is required by law when a hospital is not able to guarantee surgery within the established TMRG time limits. Thus, a patient is eligible for this kind of transfer when:

– Having an associated clinical priority level of one, 75% of TMRG has already passed without the surgery being scheduled;

– Having an associated clinical priority level of two, 50% of TMRG has already passed without the surgery being scheduled;

– Having an associated clinical priority level of three, in case of patients' request and when 5 days have passed without the surgery being scheduled;

– At 100% of TMRG, if the hospital has still not scheduled the surgery or, even if already scheduled, it has still not been performed;

– Patient's episode is not classified as untransferable (or, even if it is, but the patient requires a change of his/her status in order to be transferred);

– Patient's process is not pendent or suspended;

– There are no administrative or medical irregularities detected in the process.

All episodes fulfilling the above criteria identified on SIGLIC should, therefore, be moved to UGA's jurisdiction, so that they can be evaluated to check if they gather all required conditions for transfer. If so, TN/SVs should be issued, initiating the transferring process. In cases where transfer is not an option, the hospital should be alerted to schedule those surgeries before the TMRG.

A TN or SV is a document sent to a transferrable patient, containing a list of possible receiving hospitals for that patient to choose from. As explained in SIGIC's regulation, while a TN only lists SNS hospitals, an SV also includes private and social providers which have established agreements with the SNS. These lists are obtained automatically, through an algorithm that follows SIGIC's guidelines. All selected hospitals must have enough capacity to perform the concerning surgery, as well as guarantee all needed perioperative requirements. Hospitals on the list are ordered according to geographical proximity, to minimize transfer's inconvenience for the patient.

SIGLIC initially searches for SNS hospitals: at first, in patients' residence municipality, followed by hospitals from neighbouring municipalities and, finally, from their corresponding district. In case these SNS facilities are available to receive more patients, UGA emits a TN, through SIGLIC, addressed to the patient. On the other hand, if these SNS hospitals do not have the capacity to welcome that patient, or if waiting time already reached 100% of TMRG, UGA must issue an SV for the patient to use in any healthcare provider from a list containing public, private and social facilities.

Both TNs and SVs are sent to patients by registered mail. They are pre-numbered and untransmissible. Waiting time count is paused during the period between the emission and the utilization or refusal of the transfer offer, which must happen before the expiration date written on the document. They can only be used for the specific surgical treatment indicated in patients' care plan.

As shown in Figure 2, after a TN/SV is sent to the patient, three different scenarios can occur: the patient accepts the transfer, by directly contacting the chosen hospital from the list of options; the patient refuses the transfer and informs URGA (by mail or email) of that decision; or, alternatively, the patient does not respond. The latter situation, where the patient does not use or refuse the TN/SV before its expiration date, causes the episode to return to the transferring hospital, where the process was initially registered, which must then cancel that patient's enrolment on the waiting list. Therefore, an absence of response leads to a complete exclusion from the surgery waiting list and patients can only be readmitted if they present a written document to ULGA or URGA exposing the motives that caused the absence of response within the fixed time limit. If those motives are considered reasonable, the case can be readmitted and a new transfer document can be issued. If not, ULGA or URGA should notify the patient of their decision to not re-emit a new transfer document. If the patient still intends to be readmitted on the initial hospital's waiting list to undergo surgical treatment, then the hospital can re-enrol him/her, but waiting time is restarted and the previous position on the list is lost.

When patients refuse their TN/SV, they can also send a plausible justification to URGA. Depending on the presented motives, different actions can be adopted regarding their clinical process. For instance, if patients do not want to undergo surgery anymore or if that surgery has already been performed (either at emergency care or at a private hospital), URGA must inform the transferring hospital to cancel their

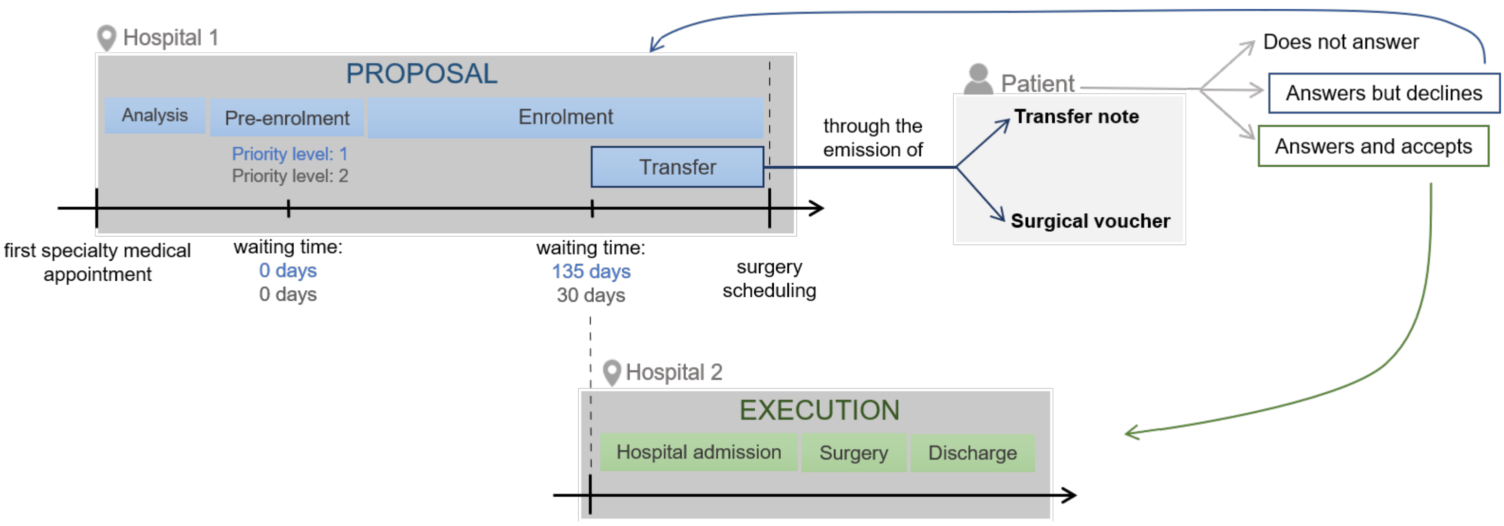

Figure 2: Representation of patients' transfer process [4].

enrolment on the waiting list. On the other hand, if patients refuse their transfer because they do not want to change hospitals, they will simply remain on their initial hospital's waiting list, in the same position as before. Once 100% of TMRG is reached, a new TN/SV is issued to them. In case of another refusal, a third SV will only be emitted if patients request so, and only if thirty days have passed since the last SV's expiration date. If patients receive a TN where there are not considered any SNS care providers near their residence area, they are entitled to decline that transfer and an SV will immediately be sent to them, so they can choose a closer private hospital. Finally, if patients invoke personal or medical motives that make it impossible for them to undergo surgery, they should ask the transferring hospital to classify their case as pendent on SIGLIC for the time being, so that no more TN/SVs are issued to them.

In case patients intend to accept the transfer, they should first get informed about all hospitals on the list of options. Once a decision is made, patients should directly contact the hospital they have chosen. Based on TN/SV's identification number, the receiving hospital must declare patients' transfer as accepted on SIGLIC and a certificate of that enrolment should be sent to patients' address, along with information regarding their rights and duties at this new healthcare facility.

When patients accept a transfer to another hospital referenced by a TN/SV, they should not have to support any additional expenses, namely regarding transportation. That only happens if patients choose hospitals which were not part of the TN/SV's options or if it is a hospital outside their residence region when there were closer options available. When no special conditions imply the use of an ambulance, public transportation is the usual choice. However, patients can opt to use their own private transport and, in that case, they have to support the corresponding expenses. According to SIGIC's rules, while the transferring hospital is responsible for transportation expenses at an initial phase (for pre-operative events and surgery), after the surgery that becomes a responsibility of the receiving hospital.

Execution, Follow-up and Conclusion

Execution includes all management processes directly related to surgery performance – from hospital admission to discharge, comprising all clinical events that occurred in between. Once surgery is performed, patients must finally leave the waiting list. If any complications develop during surgery or post-operative hospitalisation, they must be registered and treated by the hospital responsible for the surgery. When the surgeon considers patients' condition sufficiently stabilised, an hospital discharge is filled, enabling the patient to go home or to another healthcare facility.

The follow-up stage consists of all clinical events after hospital discharge and before the conclusion of an episode of care, along which patients' post-operative evolution is monitored. Regarding transferred patients, follow-up can be a responsibility of either the transferring or the receiving hospital, depending on the scope of the transfer process.

Finally, conclusion defines the moment of closure of an episode of care. If the process occurred as expected and the episode is considered completed (i.e., all planned key clinical events were performed), conclusion succeeds the follow-up stage. Nevertheless, sometimes conclusion happens prematurely, as the process might be cancelled, due to a variety of reasons.

2.3 SIGIC's impact

During the first few years of SIGIC's action, considerable improvements were attained by this programme, as it is possible to observe in Table 2. Between 2006 and 2012, despite the continuous increase in the number of patients entering the waiting list (a 38.1% growth over those six years), the total number of waiting patients experienced a decrease of 24.6%. Moreover, the median time a patient waited for surgery changed from 6.9 months, in 2006, to 3.0 months, six years later. The percentage of waiting patients who already surpassed TMRG also improved, from 43.5% (2006) to 15.1% (2012) – this indicator experienced the biggest improvement of all, with a reduction of 65.3% [17].

In fact, the continuous increase of demand did not translate into longer waiting lists due to an also significant growth in surgical services' activity. Public hospitals performed 39.3% more surgeries in 2012, when compared to 2006. Even higher increases happened in Public-Private Partnerships (PPPs) and private hospitals with agreements with the SNS to treat patients transferred from public facilities – 211.1% and 94%, respectively. The latter is also a result of the establishment of more contracts with private providers to treat SNS patients who were not timely treated at their transferring hospital. Therefore, the total number of surgeries performed reached a 54.8% increase between 2006 and 2012.

Table 2: Statistics of surgical demand and supply in Portugal, between 2006 and 2012 [17].

	Indicator	2006	2007	2008	2009	2010	2011	2012	Δ 2006/2012
DEMAND	Nb. of patients entering the SWL	451 942	517 672	534 344	566 878	573 527	611 535	624 226	38.1%
	Nb. of patients enrolled on the SWL	221 208	197 150	174 179	164 751	162 211	180 356	166 798	-24.6%
	Median waiting time (in months)	6.9	4.4	3.7	3.4	3.1	3.3	3.0	-56.5%
	Percentage of enrolled patients surpassing TMRG	43.5%	24.2%	21.6%	18.4%	13.0%	15.8%	15.1%	-65.3%
SUPPLY	Nb. of patients whose surgery was performed in a public hospital (without PPPs)	319.487	362.977	412.112	438.999	439.259	434.230	445.000	39.3%
	Nb. of patients whose surgery was performed in a PPP	11.992	12.441	13.895	12.375	19.238	26.559	37.302	211.1%
	Nb. of patients whose surgery was performed in a *hospital convencionado*	13.842	27.643	29.496	23.919	25.568	24.654	26.852	94.0%
	Nb. of patients whose surgery was performed in a *hospital protocolado*	-	-	-	-	-	18.476	25.261	-
	Total nb. of patients whose surgery was performed	345.321	403.061	455.503	475.293	484.065	503.919	534.415	54.8%

In the following years, demand for surgical treatments continued to grow, this time at a slower rate, as it is shown in Table 3. Between 2013 and 2019 it was registered an increase of 12.4% in the number of patients entering the national waiting list. Contrary to what had happened before, the actual number of patients on the list got higher during that period – a growth of 37.9%, when comparing 2013 to 2019. Response times also experienced some deterioration over the years – by 2019, 3.5 months was the median time a patient had to wait for surgery, slightly over the 3.0 months registered in 2012. Furthermore, 32.1% of patients were already waiting for a period longer than the defined TMRG (more than double the percentage of 2012) [18].

Surgery waiting list's statistics referring to 2020, although already made available by the Ministry of Health, were not considered for this research, since they were tremendously affected by the COVID-19 pandemic and, therefore, do not reflect the dynamics of the waiting list under normal circumstances, which is the object of study in this dissertation.

Overall, these indicators reflect a deterioration of SNS patients' access to healthcare. In particular, there seems to be increasing difficulties to provide timely surgical care to citizens in need of that kind of procedure. In each year, as the number of surgeries performed is always lower than the number of new patients on the waiting list, a deficit of response capacity gets accumulated. As healthcare demand continues to increase, it is crucial for the healthcare system to assure a suitable and efficient delivery capacity, following that growth.

Table 3: Statistics of surgical demand and supply in Portugal, between 2013 and 2019 [18].

	Indicator	2013	2014	2015	2016	2017	2018	2019	Δ 2013/2019
DEMAND	Nb. of patients entering the SWL	644 178	649 386	662 642	670 913	699 132	706 103	724 234	12.4%
	Nb. of patients enrolled on the SWL	176 129	184 077	197 401	210 906	231 250	244 501	242 949	37.9%
	Median waiting time (in months)	2.8	3	3.1	3.3	3.6	3.5	3.5	25.0%
	Percentage of enrolled patients surpassing TMRG	26.4%	27.8%	28.7%	28.4%	32.3%	30.0%	32.1%	21.6%
SUPPLY	Nb. of patients whose surgery was performed in a public hospital (without PPPs)	454 535	452 345	459 437	471 347	478 961	469 986	528 780	16.3%
	Nb. of patients whose surgery was performed in a PPP	47 187	52 710	53 768	53 581	55 584	59 772	33 163	-29.7%
	Nb. of patients whose surgery was performed in a *hospital convencionado*	15 915	18 336	20 054	16 200	24 608	30 962	28 204	77.2%
	Nb. of patients whose surgery was performed in a *hospital protocolado*	26 740	26 596	27 142	27 637	29 660	34 258	38 135	42.6%
	Total nb. of patients whose surgery was performed	544 377	549 987	560 401	568 765	588 813	594 978	628 282	15.4%

Since 2013, the number of TN/SVs issued to patients followed a generally increasing trend. However, TN/SVs' levels of acceptance have never reached satisfactory results, as one can observe in Figure 3. For instance, in 2019, only 18.8% of these transfer documents have been accepted by their corresponding patients [18].

The majority of proposed transfers end up being cancelled, due to a variety of reasons – including patients' refusal, TN/SVs' expiration, and surgery scheduling at the transferring hospital. Figure 4 shows the percentages of those cancellation motives over the years. Transfer refusals make up for the majority of cancellations (67.7% in 2019) [18]. Despite an increasing contribution of private hospitals to the national surgical delivery capacity, TN/SVs' results seem to fall short to reach the expected results.

Therefore, in order to guarantee a timely provision of healthcare services, it is crucial to perform comprehensive studies regarding SIGIC methodologies and their current efficacy. More specifically, it

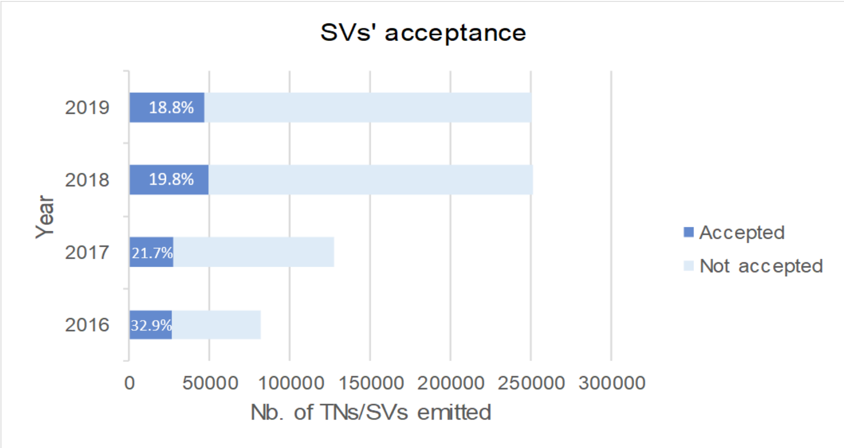

Figure 3: Percentage of TN/SVs acceptance and total number of transfer documents emissions between 2016 and 2019 [18].

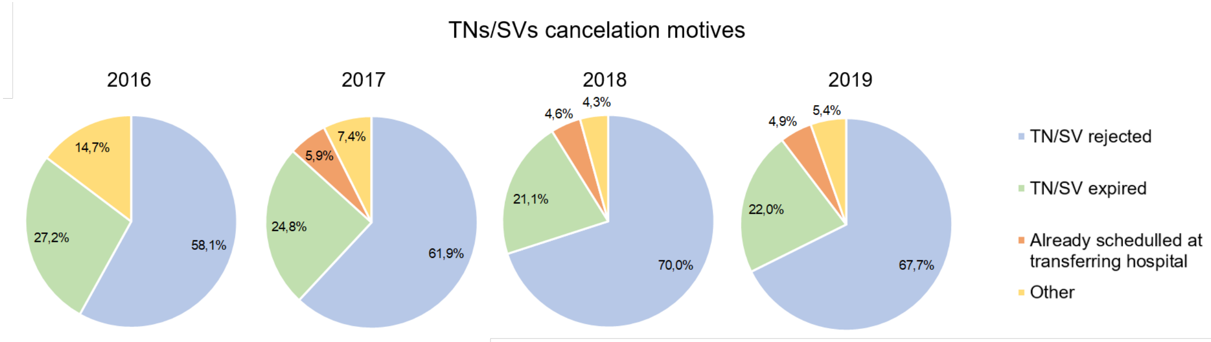

Figure 4: Percentages of TN/SVs cancellation motives between 2016 and 2019: TN/SV rejected (in blue), TN/SV expired (in green), scheduled at the transferring hospital (in orange), other (in yellow) [18].

should be of great value to understand the reasons why such a high percentage of TN/SVs end up cancelled by patients' refusal. Although this transfer system was created with the intent of protecting citizens' right to treatment within acceptable time limits, the amount of transfer rejections should be interpreted as a sign of patients' dissatisfaction with this approach.

The first step in order to study why TN/SVs' are failing to attain patients' validation is to better-understand the motives behind their decision to refuse a transfer, which would allow their health problem to be resolved sooner. A study based on a telephonic enquiry made to patients from the five Portuguese ARSs who declined being transferred [19], back in 2007, determined the main motives behind TN/SVs' refusals. These were grouped in the following four categories:

- Patients do not want to be treated by a different doctor or at a different hospital: A trusting relationship with both the doctor who initially proposed surgery and the rest of the hospital team is the basis for this type of refusal, which appears to be more frequent in older patients. Uncertainty regarding the resources available at other hospitals may also contribute to this decision.

- Patients' unavailability to use the TN/SV within its expiration date: This is usually due to personal reasons – for instance, dependent others who they need to take care of –, labour reasons – impossibility of being absent from work –, or a change of perception regarding the need for surgery – patients might feel some health improvements and refuse their TN/SVs to delay the surgery.

– Patients do not want to be treated outside their residence area: Especially for elderly patients, long distances can become an obstacle to transferring. Besides, patients often do not get informed regarding transportation offered by the transferring hospital or they might not find it satisfactory. Being away from their families and typical social environment also plays an important role.
– Lack of information: Considers all reasons linked to a lack of information, either by the patient or in the information system. Some patients showed difficulties interpreting the note sent with the TN/SV. Many times, health professionals at the transferring hospital were also not a reliable source of information, giving patients unclear or incorrect instructions.

Overall results showed that “not wanting to be treated by a different doctor or at a different hospital” was the number one given reason to justify TN/SVs refusal, with 34% of the answers, followed by “patients’ unavailability to accept the transfer before expiration date” (30%) and “not wanting to leave their residence area” (26%). Finally, a yet considerable amount of refusals (10%) were mostly due to lack of appropriate information regarding the transfer process.

An important difference was observed for patients from ARS Centre, where the geographical dispersion of healthcare facilities caused “not wanting to leave their residence area” to be the most enumerated reason for TN/SVs rejection. Patients from ARS of Lisbon and Tagus Valley, mentioned this motive much less often than patients from other regions, since in that area there is a bigger offer of both public and private care providers to choose from.

Even though this study may give a proper insight regarding the reasons behind patients’ transfer refusals, it is significantly outdated. In the initial three-year period of SIGIC’s implementation (2004-2007), to which this evaluation refers, only 17% of the total number of TN/SVs issued were declined, as stated by the same report. Over the years, that percentage has changed, reaching, for instance, 32.9% in 2016, as it is shown in Figure 3. Therefore, it is imperative to conduct further research regarding patients’ refusal motives, in order to better understand why the implemented transferring system is not accomplishing the expected results.

2.4 Conclusions

The Portuguese healthcare system has struggled to efficiently manage the national surgery waiting list for years. The legislation currently in place predicts the emission of an SV to patients whose waiting time has surpassed their TMRG, offering a list of other hospitals to where they can be transferred to in order to be treated shortly. However, the majority of patients (81.2% in 2019, as displayed in Figure 3) does not accept such proposal and prefers to continue waiting at their home hospital.

Although a study from 2007 had identified the main motives patients used to justify their transfer refusals, this subject still needs to be further explored, namely regarding what factors determine patients’ decision-making regarding their SVs. Such information would also be valuable for health authorities when developing new policies, so that they know which interventions can have an effective impact on patients’ acceptance behaviour, therefore contributing to reduce waiting times.

Chapter 3

Literature Review

This chapter explores the importance of studying behaviour, especially in healthcare contexts, and the techniques usually employed to do so. Results of previous studies related to this type of research are also presented and discussed, due to their important inspirational role in the development of this dissertation's methodology.

3.1 Behavioural Operational Research applied to healthcare

According to the Association of European Operational Research Societies (EURO), Operational Research (OR) can be defined as "a scientific approach to the solution of problems in the management of complex systems. OR can use advanced quantitative methods, modelling, problem structuring, simulation and other analytical techniques to examine assumptions, facilitate an in-depth understanding and decide on practical action" [20].

The healthcare domain is made of numerous complex systems, which may, therefore, be subjected to OR studies and interventions. OR and healthcare have been associated since the 1950s [21], mainly through attempts to optimize processes and to make care delivery as efficient as possible, searching for solutions that enable the execution of a high-quality service for patients, with limited resources and budgets spent by providers. Varying from strategic to operational planning problems, examples of OR applications in healthcare include demand forecasting, capacity planning, patient and resource scheduling (nurses, physicians, operating rooms), and treatment planning, among others [22].

Healthcare systems' functioning, alike many other areas, is highly dependent on the actions performed by the people involved in them. Regarding healthcare specifically, there are three main stakeholders that need to be considered – patients, practitioners, and health services' managers –, and whose behaviours, either on the receiving or the delivering side of that system, often affect OR analytical models developed by specialists. Therefore, models that do not incorporate human behaviour and its consequences have demonstrated to be of little practical value and usually fail to attain their expected results [23]. Besides the actions of stakeholders directly involved in the process, it is also important to mention how OR specialists themselves (modellers, analysts, consultants) and their interaction with the chosen methods to be applied in a specific situation can influence models' outcomes [24]. Thus, behaviour (in its different forms and originated from multiple sources) is intrinsically related to OR. Proof of that is the existence of an OR subdiscipline specialized in studying the connections of human behaviour and models, which is called Behavioural Operational Research (BOR).

BOR has been defined as “the study of behavioural aspects related to the use of OR methods in modelling, problem-solving and decision support” [25]. It is still one of the least explored areas of OR – a review of papers published under the scope of EURO’s working group “Operational Research Applied to Health Services” showed that approximately only 5% of those studies were related to patients’ behavioural characteristics [26]. Kunc et. al. [23] subdivided BOR into three different areas, classified according to the kind of interaction between human behaviour and OR models being studied: *behaviour in models*, *behaviour with models*, and *behaviour beyond models*.

Behaviour in models refers to incorporating human behaviour in OR models – instead of considering human beings as totally rational and even similar entities, individuals are characterized by their individual beliefs, feelings, intuitions, personal biases and backgrounds, for example. In this field, contributions from other research areas, such as psychology, are crucial for a representation of human behaviour as accurate as possible to be integrated in the OR model. This approach can be applied to study the behaviour of different stakeholders – for instance, Brailsford et. al. [27] developed a model which incorporated diabetic patients’ health-related behaviour factors to predict compliance to a screening programme, while Fugener et. al. [28] explored surgeons’ behavioural traits which influence their correct or incorrect prediction of surgery time, therefore contributing to under- or overutilization of hospital operating rooms.

Behaviour with models corresponds to how people use and interact with OR models, namely, how decision-makers select and process information given by an OR study. Heuristics and other psychological characteristics may impact the way individuals interpret and use the knowledge extracted from models. Regarding stakeholders’ acceptance and implementation of OR models’ outcomes, studies have proven that better results are achieved when OR specialists decide to use mixed methodologies instead of applying one single tool to tackle a specific issue. That happened with Sachdeva et al. [29], who analysed patient flow delays at a paediatric intensive care unit, starting with a simulation model developed with active stakeholders’ participation in order to understand what possible measures could be employed to minimize the problem. However, that seemed to not be enough to induce actual operational changes, that only occurred after other OR techniques, such as cognitive mapping, were used to validate simulation’s previous results and to identify other possible problems.

Behaviour beyond models denotes the impact of OR methods in the collective behaviour of the institution where they were applied, whether it refers to social interaction changes (collective discussion of ideas, interpretation and integration of different perspectives), or to the implementation of sustained new actions as a result of the learning process that occurred. For example, Royston et al. [30] described how the Department of Health in England achieved lasting operational improvements by using a simulation method to tackle several problematic areas of their healthcare services over a period of five years. They emphasized how sustained change can only be attained when stakeholders are brought together and included in the OR studies.

3.1.1 *Behaviour in models* applied to healthcare

Behaviour in models emerges as a crucial stream of BOR for research in healthcare, since stakeholders' individual behaviour tremendously impacts various public health policies and clinical processes. Therefore, a further understanding of health-related behaviours, namely regarding patients, and their incorporation in OR models are extremely valuable for both health authorities and medical teams. Concerning personal medical treatments, their efficiency and results are intrinsically dependent on patients' compliance with the process – for example, if patients stop taking their medication before the limit prescribed by their doctor or if they stop showing up for screening exams for a certain disease, the efficacy of such care plans is jeopardized. Furthermore, public health authorities often try to develop new policies in order to minimize harmful generalized health habits (such as smoking, sedentarism, poor diets or unprotected sexual intercourse) and to control a variety of diseases commonly associated to them (lung cancer, obesity, AIDS). Human health-related behaviour models have proven to be crucial to identify the key determinants of such dangerous behaviours and to help health authorities developing suitable policies to efficiently tackle these public health problems [31, 32].

As health-related behaviour can be a quite vague concept, it is important to better define it – it designates "overt behavioural patterns, actions and habits that relate to health maintenance, restoration and improvement", according to the Handbook of Health Behaviour Research [31]. However, not all health-related behaviours are linked to positive effects and, therefore, it is common for these habits to be divided into health impairing and health enhancing behaviours, depending on their outcomes [32]. Smoking, diet, physical activity and alcohol consumption – known as self-directed health behaviours – are among the most studied habits [31]. Sexual behaviours are also a big concern and have been targeted by public health authorities. Other areas of study include compliance with medical regimens (for example, diabetic) and healthcare services' consumption (attendances to medical appointments, screenings and vaccination sessions) [32].

By studying the literature available regarding the application of *behaviour in models* to healthcare settings, it is possible to observe common trends and, therefore, to group these articles in three subcategories, according to their main focus of analysis: modelling individual behaviour, simulating individual behaviour and, finally, a combination of both.

The first subcategory includes publications where patients' (or other stakeholders') behaviour is at the core of concern, either because they are not engaging in a specific health programme, not sticking with prescribed treatment or adopting dangerous habits, to name a few examples. The aim of these studies is to understand what are the determinants of a certain behaviour and the strength of their influence on someone's actions. Modelling the problem at hand is, therefore, a vital step in order to truly comprehend it and, if that is the case, to realize what would be the most suitable way to intervene and hopefully solve it. Thus, some papers are mainly (or only) focused on modelling a specific health-related behaviour – dissecting it in detail, along with its contributing factors (which may be, for instance, psychologic, social or demographic), the complexity of the correlations among them, and their cumulative effects –, as an attempt to formally explain usually irrational or unreasonable human actions concerning health. Subchapter 3.2. will explore more in-depth why modelling individual health-related behaviour is important, some methods used to execute it, as well as examples found in the literature of articles where modelling was the major OR technique applied to study health behaviour.

On the other hand, there are other types of studies, which are mostly directed at simulating individual behaviour. In fact, simulation is a very important OR tool: while models are crucial for deeply understanding a specific topic of concern, simulation techniques allow researchers to go even further and to evaluate different hypothetical scenarios, without having to implement them in real-life following a trial and error approach. This also reflects a shift towards a more evidence-based decision-making in the healthcare context [28]. To conduct a simulation, it is also required to develop its corresponding model; however, in the majority of the cases analysed, these models tend to incorporate very few dimensions of behaviour, since usually there is a more general focus to investigate the functioning of a larger system, where individual subjects' behaviour could have an influence on a certain outcome, but it is still not a central aspect of the study. In other words, most simulations aim at developing a virtual population of individual entities, with the capacity of showing heterogeneous behaviour that may vary according to a limited set of personal characteristics, and to study how a certain system responds to that heterogeneity of individual paths, instead of simulating behaviour per se, along with its main causes.

These more general simulations are commonly associated with the study of healthcare delivery capacity and treatment effectiveness: for instance, a group of individuals with a certain disease that regularly interact with a healthcare system unit which provides treatment for their condition. Therefore, the simulation respects to the consequences of different (and often simple) individual behaviours in that system's functioning and effectiveness, and almost completely disregards the causes and origins of those same behaviours. Furthermore, individual entities tend to incorporate more physiological than psychological variables, and their personal status tends to be essentially correlated to the individual manifestation and progression of a certain disease, for example, and not so much linked to health-related decision-making. This kind of articles, which only slightly incorporate behavioural factors as part of a larger simulation model will be further scrutinised in Subchapter 3.3.

The third and final category, which will be analysed in more detail in Subchapter 3.4, refers to the combination of both a meaningful modelling of the health behaviour being tackled (including its contributing factors, as well as the strength of their correlations) and a simulation, which incorporates all those key determinants and allows to examine their role in hypothetical behavioural changes. Such studies present a more complete analysis and are of tremendous value for a better understanding of health-related individual behaviour, thus matching the purpose of this dissertation.

3.2 Modelling individual health behaviour

Modelling a behavioural problem might be a necessary step to plan an intervention or, alternatively, it can be seen as an end in itself, since it promotes a better understanding of any given situation. Nevertheless, it shall not be seen as an easy task, since it is essentially an attempt to formally (and sometimes mathematically) structure human actions and decisions, which are often known to lack logic and rational bases. Despite being a difficult process, efforts have been made to improve the art of developing such models, especially when public health policies' results are at stake.

Different modelling theories have been applied in the literature over the years, without attaining a consensus on what are the most suitable for structuring individual health-related behaviour. As every method has its own pitfalls, which will be further discussed in this chapter, modellers are forced to make concessions along the process, while trying to not significantly compromise the reliability of their work. Due to the lack of clarity that surrounds heath behaviour studies, a lot of theoretical models are mostly based on assumptions rather than evidence, which also hinders their application to real-life scenarios.

To incorporate behaviour into OR models it is important to comprehend the factors that determine a certain action and how they might be intertwined. Since such connections are usually far from being simple, this demands for an exhaustive information gathering that can be executed using various methods (directly – communicating with key players through interviews, surveys, or meetings –, or indirectly – resorting to comprehensive literature reviews), and that can come from different sources (stakeholders, experts, researchers) [33]. It is also important to keep in mind that, although every model is ultimately a simplification of reality, efforts should be made to incorporate all the requisite factors to correctly portray the problem at hand, in order to guarantee the reliability of models' outputs.

3.2.1 Stakeholders' consultation

By modelling a certain problem, we are trying to define it, as well as all its components and determinants. This detailed examination facilitates a better understanding of the situation of concern, and it may even contribute to find a possible solution. OR researchers are experts who master numerous techniques for modelling different situations; however, most times they are not familiarized with the context of the problem they are going to help tackling. Although the intervention of an outsider can be beneficial to help examining a problem from a new and impartial perspective, that can also mean that important features go unnoticed and end up not being considered in the developed model. That is why stakeholders' consultation is so important to this process [34].

If one wants to comprehend a problem and its contributing factors, it is vital to listen to the people who are actually experiencing it – their perspectives, perceptions, tacit knowledge, and opinions – instead of simply making assumptions regarding their points of view. Taking those inputs into account will enable a more accurate and complete model design. Furthermore, it will also increase trust and create a feeling of involvement in the project, which may also promote acceptance of its results. This is of even greater relevance when dealing with behavioural aspects. By involving stakeholders during the process, researchers will be able to examine their behaviours regarding the problem, the reasons behind those specific actions, and to identify opportunities to positively influence them. In the end, it does not only contribute to a better representation of the problem itself, but also increases the success chances of a possible intervention, since people will be more prone to participate and, if that is the case, to modify their usual behaviours [34]. In fact, research has proven that the less involved stakeholders feel during the modelling stage, the less credibility they attribute to that model's outputs and the less receptive they feel to the implementation of hypothetical changes. That is why healthcare OR in general (and not only when behaviour is the main focus of interest) has been investing in the potential value of patient and public involvement (PPI) in research [35]. The same is true for other stakeholders besides patients. For instance, when dealing with healthcare practitioners, their participation in the OR process is also crucial for their posterior receptivity and acceptance of the obtained results [36].

Although stakeholders' consultation is important in each step of the way, it is decisive especially in the initial stages, to help defining and understanding the problem. Nonetheless, stakeholders can also be very useful at a later stage, to validate the developed model [34].

3.2.2 Psychological models of health behaviour

After the information gathering stage, it comes the time to combine and structure those findings. Psychological models specially developed to represent human health-related behaviour are a useful tool to do so. Two of the best known and most commonly used models will be explored next.

Health Belief Model (HBM):

Initially developed by Rosenstock in 1974, the Health Belief Model is one of the oldest and most used psychological models to accurately explain individuals' health behaviour. According to HBM, the likelihood of an individual taking a certain health-related action or adopting a given behaviour is related to one's perceptions regarding their own health status, the implications of a subsequent health problem or disease, and the pros and cons of following the behaviour in question. These are translated into four components of the model: perceived susceptibility (perceived individual risk of contracting a disease/developing a condition) and perceived severity (perception of the harmfulness of a disease and its consequences), that together make up the "threat perception" dimension of the model; perceived benefits (personal assessment of the benefits of adopting the behaviour to avoid a disease) and perceived barriers (evaluation of the disadvantages of adopting the behaviour), which represent the "behavioural evaluation" part of the model. Therefore, a person engages in a certain health-related action when the net benefits of taking action supplement the threat of not adopting the behaviour [27, 32, 37].

Besides these, the model comprehends two other components: health motivation and cues to action. The former reflects someone's general interest and concern regarding personal health, while the latter represents triggers that incentivize the adoption of the behaviour and that can be either internal – a physiological condition (the appearance or worsening of physical symptoms, for example) – or external – mass media campaigns, a health professional advice or another type of social interaction [27, 32, 37]. The structure of the model can be observed in Figure 5.

Recently, self-efficacy, a construct initially employed by Bandura's Social Cognitive Theory (SCT), has been added to HBM to represent someone's perceived capacity to carry out certain actions [37].

The biggest advantage of HBM is its simplicity, i.e., the variables are quite straightforward and simple to understand, even for someone who is not familiarized with psychological terms. That made it easier for this model to overcome the barriers of psychology and to be applied by other disciplines, like OR. Nevertheless, and despite having been widely used in different areas, it lacks some structure, as the mathematical relationships between variables are not well-defined (it is usually assumed an additive relation between them, but it is not formally stated). In addition, there is the fact that this model largely relies on constructs related to individual perceptions (of susceptibility and severity, for instance), which can be very difficult to quantify due to their subjective nature [27, 38].

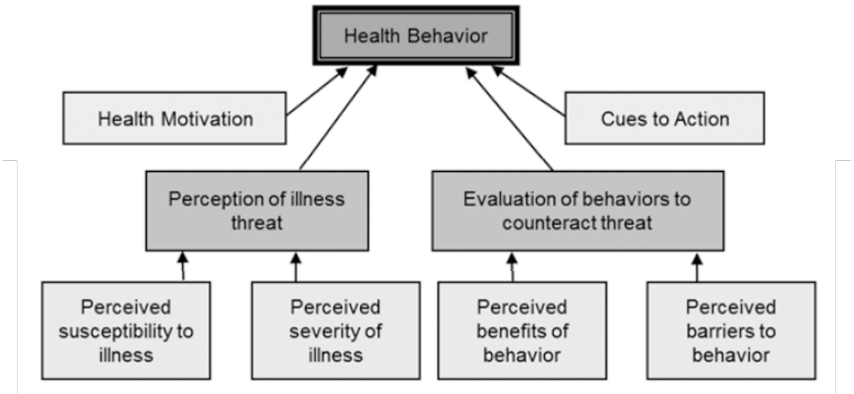

Figure 5: Health Belief Model [38].

Theory of Planned Behaviour (TPB):

The Theory of Planned Behaviour was developed by Ajzen as an extension of the Theory of Reasoned Action. The corresponding model shows that one's intention to follow a health behaviour depends on three factors: attitude towards behaviour, subjective norm, and perceived behavioural control. Attitude towards behaviour is determined by personal beliefs regarding its possible outcomes and their corresponding evaluation. This construct translates if a person values or not the effects of the considered behaviour. Subjective norm reflects the extent to which people whose opinion one values want them to engage in a certain behaviour, and it is influenced by normative beliefs and motivation to comply with such behaviour. Lastly, perceived behavioural control, as the name suggests, regards the degree to which people consider they have control over the attained results, i.e., if one has the necessary resources and/or capacity to execute a certain action. Perceived likelihood of occurrence and perceived facilitating/inhibiting power are the two variables that influence this third dimension of the TPB model. Affecting all aspects of the model, there are external variables, which can either be demographic or personality features [27, 32]. Figure 3 shows TPB model's structure.

This model presents a well-defined mathematical structure between the different variables, making it a good candidate to be implemented in practice. A drawback of this theory is related to the fact that it was not specifically developed to be applied to health-related behaviours and so, it neglects the health threat dimension of the problem, which is a major determinant of this kind of behaviour and is present in other models, like the HBM [27].

Although there are other theories beyond the ones described above, these were the psychological models of human behaviour that were found during this literature review to be more suitable to be applied to healthcare contexts similar to the one being studied in this dissertation.

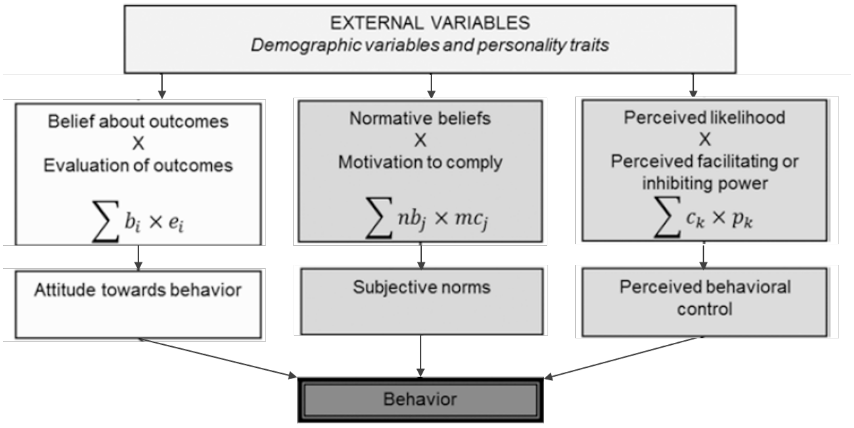

Figure 6: Theory of Planned Behaviour [38].

3.2.3 Relevant examples found in the literature

Recently, in 2020, Hsieh and Lai [37] developed an integrated model to identify the psychological determinants behind Taiwanese citizens adoption of a personal health passbook technology, designated "My Health Bank" (MHB), aimed at promoting health self-management among the general population. The passbook gives people access to their exams results, enables them to book medical appointments and to monitor their general health condition, as well as to search for preventive healthcare information. Almost four years after its implementation, only 3.2% of the population was using MHB, fact that incited researchers to find out which were the key factors behind people's intention to use (or not) the health passbook. The developed research model integrated the Technology Acceptance Model (TAM) with two complementary models of health-related behaviour: HBM and the Protection Motivation Theory (PMT). HBM and PMT overlap in some of their constructs, a fact that was taken into consideration by the researchers, who implemented a combination of both models without repeating similar concepts. Health literacy was another determinant added to the integrated model. All these constructs contributed (positively or negatively) to the chosen outcome measured: intention to use. Initially, hypotheses were made regarding the positive or negative nature of the relations displayed in the model: perceived susceptibility, perceived severity, self-efficacy, cues to action, perceived utility, perceived easiness-of-use, and health literacy were assumed positively related to intention to use; contrariwise, rewards and perceived barriers were expected to have a negative impact on people's willingness to use MHB.

The identified hypotheses were then evaluated through a questionnaire, consisting of different sets of questions, one for each construct of the model. The answers were qualitative, according to a seven-point Likert scale. Structural Equation Modelling (SEM) was the method employed to study the causality between the model's parameters, based on the questionnaire's answers. The results confirmed all the previously established hypotheses, except for the positive correlation between perceived severity and intention to use. It turns out that the mentioned construct did not have a significant impact on people's acceptance of the new technology. The authors justified it with the fact that, when facing a severe health problem, patients tend to seek face-to-face treatment at a healthcare facility, instead of resorting to a digital tool. The remaining options were validated by the survey's results, proving that health belief factors, such as perceived susceptibility, rewards, perceived barriers, self-efficacy, and cues to action, as well as health literacy and technology acceptance elements (perceived utility and perceived easiness-of-use) are all key determinants of people's behaviour regarding health technology acceptance, more specifically, of their willingness to use the health passbook.

An important contribution of this study was the integration of multiple models, mixing health-related behaviour theories with a technology acceptance model. This enabled a more accurate representation of the problem at hand, whose validity and reliability were demonstrated by a higher explained variance than the ones obtained in previous similar studies, according to the authors. By revealing not only which were the determinant factors for MHB's acceptance among patients, but also the strength of their impact, this study provided valuable information to Taiwan's health authorities, which could therefore predict what kind of policies and strategies would be more effective and actually induce behavioural changes in the population, promoting the use of the health passbook.

A different study applied the HBM model to analyse the predictive factors of oral hygiene habits among young girls in Iran [39]. Data regarding demographic aspects and the causal constructs of the psychological model employed was collected through a survey. SEM was used to identify the correlation weights between each of those factors and dental habits. The obtained results pointed perceived barriers, self-efficacy, and cues to action as the main causal aspects of the analysed behaviour. Therefore, the authors recommended that Iranian oral health programmes should try to approach these psychological dimensions with their target audience in order to attain satisfactory outcomes.

Patients' compliance with screening programmes has also been the target of some behavioural studies. Regarding diabetic retinopathy, Brailsford et al. [27] developed a model to predict patients' intention to attend an examination session for which they were previously invited, mainly based on HBM. Later on, Brailsford et al. [40] applied a TPB model, once again to predict patients' compliance, this time for breast cancer screening sessions. As both models were used not only to understand and structure those screening programmes, but also to simulate the influence of individual behaviour on their outcomes, they will be dissected in more detail ahead, in subchapter 3.4.

3.3 Simulation of individual health behaviour

Simulation techniques are commonly used when applying OR to healthcare contexts, facilitating decision-making and problem-solving by mimicking how complex systems evolve over time. By simulating a certain reality, different policies can be tested before they are truly implemented, enabling cost-effectiveness assessments and comparisons of various possible scenarios. Furthermore, it allows the performance of clinical trials which would not be feasible to conduct in real-life, either because they would imply the adoption of unethical methods or require a lifetime following of the targeted population, as it happens with the study of screening programmes' effectiveness for example [41].

Simulations are extremely useful since they enable researchers and health authorities to test reforms or new policies, predicting their expected impact. They can be used by healthcare managers, for example, to predict the cost-effectiveness of a new technology or treatment, helping them decide whether or not to invest [41]. Some simulation methods are also able to include patients' individual behaviours and can even incorporate the psychological constructs of the theoretical models described above.

In the literature, it is possible to find many applications of different simulation techniques to healthcare settings. However, it is not so common for those simulation models to incorporate individual behaviour and, even when they do so, it usually lacks the deterministic dimensions that lead to those actions. That can be unexpected, especially if one thinks about how human health-related behaviour (either from patients, practitioners or managers) has the potential to influence health strategies' success. While superficially including human behaviour may be enough in some cases, particularly when the focus of study is on a larger dimension (like an evaluation of the overall functioning of a healthcare facility, for example, where it is important to include a few individual characteristics of patients in order to simulate how they will heterogeneously use that service and consume available resources, but there is no interest in studying the causes behind those individual behaviours), the same is not true when researchers intend to investigate ways of inducing behavioural change, since it is crucial to deeply understand how certain factors influence subjects' actions in order to plan and simulate this kind of interventions.

Along this subchapter, there will only be explored simulation studies which include behavioural features as part of their models, although still not in a detailed manner, without referencing their causes and determinant factors. Thus, their main goal is not to investigate the impact of behavioural changes, despite still not completely neglecting the impact of individual decision-making and actions in the evolution of the systems under study.

The two main simulation techniques used to incorporate individual behaviour are going to be explained next. According to the literature, these are the most suitable for the task, since they admit the existence of individual units (entities or agents) which assume, at a given instant, a certain set of variables that characterize them. They can exchange messages with the environment surrounding them and take different actions accordingly to those stimuli and their personal traits. Therefore, each unit can act differently from the others, originating a heterogeneous set of "characters", appropriate to represent individual human behaviour.

Discrete-Event Simulation (DES):

DES is probably the most commonly employed simulation method, usually representing operations where individual behaviour may play a significant role (regarding healthcare, an example might be patients' flows and use of hospital units) [27]. This method is usually applied for modelling systems that can be represented by individual entities which undergo a series of succeeding discrete events. Entities have attributes and are usually inserted in an environment, with which they exchange messages. The system is regulated by different state variables, which can suffer changes over time [42].

The system needs to keep track of time, of a list of ordered forthcoming events and of the values assumed by the state variables. As time progresses, the simulation also advances: the next event on queue is processed, which can cause alterations on some state variables, that should also be recorded [42]. The relationships between state variables and entities, which ultimately determine each of their behaviours, are regulated by rules encoded in the simulation program by the OR modeller [43].

DES is commonly used to analyse systems' performances and to identify processes' bottlenecks. Although entities do not typically communicate with each other, they can still be programmed to do so, as it happens in some models of infectious diseases [42, 43]. In order to use DES to model human behaviour, entities should be given attributes and properties which correspond to human behavioural features, namely related to health [23].

Agent-Based Simulation (ABS):

ABS is a simulation technique used to model complex systems where autonomous, heterogeneous and intelligent entities (here called agents) interact both with the environment and among themselves (according to a set of predefined rules introduced by the modeller) to attain a certain goal [42, 43]. ABS is the tool of choice to simulate agents' knowledge and information processing, since agents are able to memorize past experiences and learn from them, adapting their future actions accordingly [43]. Individuals' independent behaviours are determined by their state variables at a given instant in time and can be divided into reactive (when an individual's action is a response to a stimulus or event) or deliberative (when it involves a conscious pursuit of the agent's goals).

Until now, ABS has been less frequently used for OR studies than DES. Nonetheless, it offers a promising approach to simulate human health-related behaviour due to its ability to represent autonomous units that can communicate with each other, memorize experiences, learn from them and change their individual actions in an adaptative response, therefore, suitably mimicking human beings [23]. Despite all this potential to be applied in healthcare contexts, some researchers still argue that, as far as BOR is concerned, DES is sufficiently advanced to similarly simulate patients' health behaviours, with the advantage of being allegedly simpler to develop and perform, when compared to ABS [43].

Besides DES and ABS, there are other simulation techniques that, although not explicitly developed to incorporate human behaviour, can also be used together with the aforementioned methods in larger system simulations. System Dynamics (SD) is one of those examples, a technique that depicts systems from a more general point-of-view and that is commonly used to identify causal relationships and linkages among variables. It is usually applied to design and test interventions or policies [42].

Hybrid models, which integrate more than one simulation method, can be especially useful in more complex studies, where the overall problem is made of different subsystems, which require distinct modelling or simulation approaches and that cannot be studied in isolation. Despite its advantages, using hybrid models implies a significantly higher complexity during the development stage, as it might require the integration of different software [43].

3.3.1 Relevant examples found in the literature

The use of simulation in healthcare settings is abundant in the literature. In fact, over 40% of BOR papers categorized as incorporating behaviour in models applied simulation as the major OR technique of the study [23]. The following examples show some of the potential of considering behaviour as part of a larger model, typically representative of the operational functioning of healthcare services, and represent a starting point for further and more complete behaviour analyses that could be performed in the future.

Lopes et al. [44] developed a simulation model to forecast the evolution of the medical workforce until 2050, in Portugal. Both supply and demand of physicians were taken into account in a robust model intended to help healthcare services planning. ABS was the chosen simulation approach, due to its capacity to account for individual preferences and decision-making, which play an essential part in the context of this study. According to the authors' best knowledge, that was the first time ABS was used to simulate future medical workforce demand.

Corresponding to the supply side of the simulation, each model's agent represented a physician that evolved in the system over time, passing through a sequence of stages that represent their lifecycle: from birth to medical school, including intern practice, speciality school and graduation; then, the process of joining physicians' active workforce and, finally, reaching retirement. Through all of that, agents' personal choices and behaviour influence their path in the system: they can drop out of medical school, perhaps they might need more than the minimal number of years to complete school, they may or may not obtain a speciality vacancy and that might require them to repeat the admission exam, they are also able to opt between practicing at the public or private healthcare sector (or both) and, finally, they might emigrate or die at any moment of the simulation.

In order to properly model all these uncertain steps, which are associated with each subject's personal experience and decision-making and, therefore, vary from individual to individual (or agent to agent), the authors relied on historical trends and statistical distributions that best fit each of the model's dimensions. For instance, educational stages (medical school and speciality school) were represented by exponential functions of time, while emigration rates were inferred from the scarce information available. Specialities were ascribed randomly, while in reality they depend on students' grades on an admission exam. Nevertheless, real speciality vacancies numbers were respected and when students were assigned a speciality with no free spots left, they would have to wait one more year to apply.

The demand part of the simulation was based on demographic projections and their influence on the overall population's quest for health services. Particularly the age distribution of the Portuguese population seems to have a significant impact on medical services demand: as people live longer, they consume more medical resources and ultimately represent a bigger burden for the healthcare system.

Overall results suggest that physicians' workforce will probably be insufficient to meet the care needs of the population. Besides modelling the problem, the authors simulated the effects of two possible public policies: raising physicians' minimum retirement age, and increasing the number of spots in pre-grad medical school and in speciality school. Both strategies resulted in a workforce increase, weakening the physicians/population imbalance.

Despite the suitable use of ABS to address this forecasting problem, due to the critical role that heterogeneous physicians' individual behaviour and decision-making play in the simulation model, the authors recognize there is still some unemployed ABS modelling potential that should be explored in future studies. For example, the model could incorporate microeconomic foundations to assess how individual economic behaviour can influence the simulation, namely regarding wage and emigration. Moreover, since ABS allows for agents to communicate and be aware of their surrounding environment and space, regional asymmetries within the country regarding physician's demand and supply is another topic that could also be studied using this same model. Therefore, there is still space for other studies to further explore individual behaviour modelling regarding practitioners' workforce planning.

Another example of the application of simulation techniques to healthcare is the article written by Lebcir et al. [45], in 2017, where a DES model was built to represent the care structure of Parkinson's Disease, in the UK. The model included patients' pathways in the healthcare system (including primary and secondary care), different diagnosis and treatment options, disease phases, and the mix of resources employed treating these patients. DES was the chosen simulation method, since it allows for patients to be represented at an individual level by entities assuming different characteristics, such as age, gender, disease progression, treatment approach, type of medical following, and the specific pathway tracked in the care service system. By monitoring entities' attributes evolution over time, it is possible to observe how the simulated advances and what changes of status occur. DES also allows for the integration of resource and capacity constraints, which are very important in healthcare studies.

The focus of this article was to assess the impact of an increased use of community services by Parkinson's Disease patients, providing some of the services typically delivered by secondary care units. Four scenarios were simulated: a 10% (low), 20% (medium), 40% (high), or 50% (very high) increase in the number of Parkinson's Disease patients treated at community care facilities. Results show that these shifts of patients from hospital to community care positively impact hospitals' budgets and, therefore, this DES model should represent a useful tool for health managers and authorities when developing new plans to restructure and optimize the Parkinson's Disease care system implemented in the UK.

Some limitations of this study include the disregard for patients' psychological characteristics which may significantly influence disease progression, attendance to scheduled appointments, and treatment efficacy. Furthermore, social status and comorbidities were also absent from the model.

One of the purposes of simulation is to conduct clinical trials that would not be feasible to do in real-life, for example due to potentially unsafe methods to the targeted subjects. Day et al. [46] resorted to a hybrid model to perform a simulated randomized controlled trial, where they compared the vision loss extent of diabetic patients for varying diabetic retinopathy's screening intervals. Thus, the authors were able to test the implications of different health policy programmes without worrying about possibly harming real patients.

Two different simulation models were integrated in this study: an ABS model representing the targeted patients (veterans who suffered from diabetes) as individual agents, and a DES model which depicted the eye clinic system. Together, they formed a hybrid model, where each agent, which was given a certain set of individual characteristics (age, duration of diabetes, body mass index, current tabaco use, etc.) derived from a real-world cohort study with diabetic veterans, interacted with the eye clinic system once it was time for a scheduled appointment. At that moment, an entity would be created in the DES model, including all the information of the agent from which it was originated. Entities that exist in the eye clinic model can undergo different procedures, such as screenings to determine their current state of diabetic retinopathy, and pan-retinal photocoagulation laser surgeries. If any changes in the health status of an entity are detected (change in diagnosis or indication for surgery, for example), that information should then be transferred to its corresponding agent in the ABS model.

Different groups of patients were created, where screening sessions were scheduled at different time intervals: yearly, for the control group, and ranging from every two to five years, for the four experimental groups, studied in 10 year-long simulations. Results, as expected and hypothesised by the authors, show that switching from annual to biannual screening appointments induces no significantly higher risk of vision loss in diabetic veterans, who do not yet suffer from diabetic retinopathy. However, if only scheduled every three years, the outcomes do change, and an increased risk of vision loss was observed. Nevertheless, as no significant difference was found when changing the screening interval from one to two years, health policies should be optimized in order to save resources, while still providing a similar level of quality of care for diabetic patients.

Similar approaches, where an ABS model that simulates the progress of a certain disease within a population of individuals is used in integration with a DES model representing how these agents compete for resources in a healthcare service environment, can be used to study how patients truly interact with health systems and how current policies may or may not be optimized. Furthermore, resorting to simulation techniques enables researchers to evaluate any possible scenario without causing patients any harm and provide useful evidence for the real-life implementation of new health reforms.

Viana et al. [47] also developed a hybrid simulation model, this turn combining DES and SD, to model Chlamydia infection in a community of the UK. Here, while the DES model represented an outpatient clinic that treats not only Chlamydia but also other sexually transmitted infections, SD depicted the progression of this disease within the studied population. To sum up, the overall model translates how operational level decisions at the clinic influence and are influenced by Chlamydia's progression among the population. The infected group of people in the SD model is converted to patient arrivals in the DES model, that then, depending on the clinic capacity and available resources, return to the SD model as recovered or still infected individuals.

The clinic was searching for new ways of structuring patients' walk-in process (since patients did not have to schedule appointments in advance nor it was required for them to be referred from another physician), so that they could properly treat the highest number of people possible, preferably without causing long queues and waiting times. Patients' pathways inside the clinic were modelled and information regarding processes' durations, patients' demographic characteristics, available resources, and staff's schedules was collected.

Two different scenarios were evaluated: one that depicted the actual resource capacity of the clinic at the time, and another one where resources were increased to maximum levels. Results showed that, for the second scenario, all three main outputs (number of patients leaving the clinic without being seen by a doctor, prevalence of Chlamydia in that population, and estimated medical costs) decreased, when compared to the first scenario. Therefore, despite falling short to incorporate heterogeneous individual behaviour, this study provided a clear insight regarding the tremendous effect that resource allocation and other management decisions at healthcare service units may have on public health (and how simulation might be helpful in order to implement successful actions).

3.4 Modelling and simulating individual health behaviour

As already stated, despite the fact that BOR has been applied to healthcare contexts multiple times over the years, few are the studies that contemplate both a detailed modelling and simulation processes of the psychological fundamentals of individual health behaviour. There are cases of such models that never get to be applied to simulations, or simulations that actually consider some kind of individuality, although usually more at the physiological level. Considering the context of this dissertation and other similar healthcare problems, it is important that researchers also turn their attention to the psychological factors that determine each decision people make regarding their own health.

Why do sick people stop taking their prescribed medication? Why do risk groups for a specific disease miss their scheduled screening exams? Why do people enrolled on very long waiting lists turn down the opportunity to receive faster treatment? It is in order to understand the answers to these questions and the reasons behind such apparently unreasonable decisions that studying psychological aspects related to health (of course, without neglecting the physiological matter which is also crucial) is so important. In the end, it is useless to develop new health policies if the people they are targeted at do not engage with them.

The best examples found in the literature that illustrate this line of thinking were both developed to evaluate compliance with screening for two distinct diseases, diabetic retinopathy and breast cancer. Their models incorporate both physiological and psychological individuality of the studied populations, considering the evolution of the disease at hand, as well as psychological models that provide theoretical support for each person's decision to comply or not with a certain screening programme. Simulations were run to test and compare the outputs of different scenarios and interventions, which proved that positive physiological outcomes (years of life saved, for example) could be achieved not only by tuning operational aspects of the screening programme, but also through other kind of public interventions that have the capacity to change the mindset of those patients, such as awareness campaigns.

In 2003, Brailsford and Schmidt [27] conducted a pioneer study where they developed a model for predicting diabetic patients' compliance with a screening programme for diabetic retinopathy. The research was based on a previous study, which modelled the impact of this programme considering a constant compliance value for all patients involved. That meant that different people supposedly presented the same probability of attending a screening visit, independently of their background, personality traits or history of previous attendances. On the other hand, Brailsford and Schmidt realized the crucial impact that individual behavioural factors played in healthcare OR systems and used the HBM to incorporate health-related behaviour into the model. Other components – physis, emotion, cognition, and status (PECS) – were embedded in the entities of a DES model. Compliance was computed as the product of: visits, which considered the history of previous attendances for screening; motivation, which could assume three different values ("low", "medium" or "high"); and a PECS estimate, that was computed as the average of its four elements. Physis was determined by the stage of diabetic retinopathy among three possible levels. Emotion was the result of the product of “anxiety” and “perceived susceptibility” (both could be “low”, “medium” or “high”). Cognition was the product of “knowledge of the disease” and “belief about disease prevalence” (“low”, “medium” or “high”). Finally, status was given by individual’s educational level (which assumed a numerical value depending on its classification as “low”, “medium” or “high”). The result of Compliance was a number between zero and one which, if higher than a uniform random number that was sampled, meant that a specific patient would attend the screening visit.

As this was essentially an exploratory study to evaluate if the model was sensitive to assumptions regarding individual behaviour, many of these mathematical expressions and values were chosen arbitrarily. The model was run for ten different scenarios. In each scenario, different components were being varied to extreme values in order to understand their effects on compliance and, consequently, on patients’ number of years of sight saved (the chosen outcome measure). The results revealed that not considering individual variations of compliance with screening leads to an overestimation of patients’ attendance and, consequently, of that programme’s beneficial effects. This represented a major novelty for healthcare OR, since it demonstrated the impact that behavioural aspects could potentially have and how, by neglecting individual behaviour, researchers were obtaining imprecise and unrealistic results.

Later, in 2011, Brailsford et al. [40] developed another simulation study, this time designed to model a breast cancer screening programme. By including behavioural aspects of the women targeted by this programme, the authors were able to identify which were the key factors that influenced attendance for mammography and to observe the extent to which it affected the success of this public health strategy. The patients’ behavioural variables considered, as well as their mathematical correlation, depended on the approach chosen to model attendance, within the four options applied: “local” and “global” attendance percentage, TPB, and Baker and Atherill’s compliance model. Local and global percentages are the simplest methodologies among the four, as they only require the definition of a parameter, p. Local percentage of attendance implies that every woman attends p% of their screening sessions, while global percentage means that p% of all women attend every scheduled screen and 1-p% never get screened. Regarding psychological factors, TPB was the chosen model due to its well-defined formal structure. The model’s variables were extracted from a questionnaire answered by 2058 women in a previous study, which comprised specific questions developed to measure TPB’s qualitative constructs, as well as some demographic and socio-economic topics. Individual answers were given on an ordinal scale and, afterwards, TPB’s constructs were computed: attitude towards mammography, subjective norms, and perceived behavioural control. These three variables were linearly combined to predict individual’s intention to attend. Intention to attend together with perceived behavioural control were used to predict ultimate behaviour (attend/not attend). Finally, Baker and Atherill’s compliance model is based on patients’ screening attendance histories regarding previous sessions. Recent attendances/non-attendances had a higher weight when it comes to predict future behaviour, rather than distant history.

The created model was based on screening and physiological parameters, besides the already mentioned behavioural factors, and was then used to perform a DES. Screening parameters are associated with the frequency and detection capacity of the screening exam. The former is defined by the health policies currently in place (or that are being studied in this situation): how often should patients get mammography screenings and at which age stamps. The latter is related with mammography's sensitivity (the capacity to detect existent tumours), which is mainly dependent on the tumour's dimensions at the time of the session. Physiological parameters were also included to describe mortality, tumour growth, age of cancer onset, and survival. Within the model, each woman lives a simulated life, from birth to death. Along the way, there is a set of events that may or may not happen: develop breast cancer, be invited for a mammography screening and attend that same exam session. In case a subject develops breast cancer, it can be detected through the screening programme or it could be self-detected. It can, then, lead to cure or death by breast cancer. Death from other causes is also considered in the model.

The model was used for comparing the outcomes of five screening scenarios (where different mammography's' frequencies and patients' ages were defined), as well as the differences of considering different behavioural and cancer growth assumptions. Results show that there are two screening scenarios that seem to attain better results when it comes to the percentage of screen-detected breast tumours than the current UK mammography screening policy (targeting women from 51 to 69 years old, every three years), which consist of either decreasing screening starting age to 45 years old or increasing screening frequency to every two years, instead of three. When running the simulation model using both the global percentage and Baker and Atherill's compliance model to predict attendance behaviour, no significant differences were found regarding the percentage of screen-detected tumours for the two most favourable screening scenarios. However, when using local percentage or TPB, screening women from 45 to 69 years old every three years displayed better outcomes. These two methods for attendance prediction were found to be the most suitable for this type of study, since they also exhibited similarly proportional results across all the screening scenarios.

TPB's formal structure and its three main constructs allowed the authors to perform a sensitivity analysis of the behavioural variables. Thus, it was possible to conclude that an increase in the number of tumours detected could not only be achieved through restructuring the screening policy in terms of targeting ages and frequency, but also through actions aimed at changing patients' psychologic parameters regarding breast cancer. For example, an increase of 10% in perceived behaviour control led to 3% more tumours screen-detected, an increase of 2.5% in the average number of attendances and of 2% in life-years saved. Changing the other two constructs (subjective norm and attitude towards behaviour) resulted in smaller variations of the outcomes. The cumulative impact of changing all three behavioural constructs at once was even higher, causing a 4% increase in the number of cancers detected, as well as in life-years saved, without modifying the scope of the mammography screening policy at all. This study led to the important conclusion that behavioural interventions can significantly impact public policies' efficacy and should also be considered by health authorities as a means to achieve better results.

On the topic of surgery vouchers' acceptance behaviour, there were not found any studies. Nevertheless, similar approaches to the ones applied to predict appointments' attendance, compliance with screening, and health technology's acceptance may be adapted to tackle that problem, as they might allow researchers to better understand the fundamental reasons behind patients' decision to decline the voucher (and subsequently, the opportunity to gain health through having earlier surgery). A comprehensive behavioural model could then give some insight regarding why the surgery vouchers' strategy is not obtaining as good results as expected and provide a tool to simulate the outcomes of patients' behavioural changes in the system currently in place.

3.5 Conclusions

As it was explained along this chapter, it is essential for researchers to include stakeholders' behaviour in their healthcare OR studies, due to the tremendous impact of their actions in services' functioning.

Since the aim of this dissertation was to conduct a *behaviour in models*' study, where patients' acceptance behaviour regarding a surgery voucher offering programme is explored, it was important to review the available literature on the subject, namely searching for recurrent themes and methods typically applied. Health impairing behaviours is one of the main themes explored, as it has a considerable negative impact on public health. In addition, health services consumption is another area of great importance which evaluates patients' attendance to scheduled appointments or screening exams, for example.

The methodology applied in such studies varies according to the main objective of analysis. Along this literature review it was possible to subcategorize them in three groups, according to observed patterns: articles that deeply explored psychological, physiological and demographic determinants of a certain behaviour and developed its corresponding causal model; articles that, despite neglecting important determinant factors that regulate individual behaviour, used simulation to study how behavioural changes could influence the functioning of a wider model; and, finally, a few articles that seemed to put it all together, with a detailed and comprehensive model of individual behaviour later applied to simulate the impact of possible interventions.

While simulation studies were abundant in the literature, they mostly missed to cover the effects of psychological traits on behaviour, since they usually explored the dynamics of a "whole system" view of health services, where individuality had an influence to some extent, but it was not required to dissect its contributing factors. However, that cannot be ignored in the case of this dissertation, where the main goal is to understand why patients make the decision to reject the opportunity to wait less time for their surgeries. Both physiological and psychological characteristics play an important role regarding health-related decision-making and, therefore, none of them should be ignored.

Regarding voucher acceptance behaviour, no relevant BOR articles were found, which constitutes a gap in the literature this dissertation intends to close. Nevertheless, the methodologies here explored, although applied to slightly different contexts, represent a significant starting point for choosing the most suitable research procedure to apply in this specific project. Namely, the last few studies covered in this chapter (about compliance with diabetic retinopathy screening [8] and breast tumour screening [24]) constituted a relevant foundation for the development of this dissertation, since they approached a health-related behavioural problem using a very comprehensive and complete OR methodology.

Chapter 4

Methodology

This dissertation encompasses a comprehensive study of the SVs' system implemented in Portugal with the aim of minimizing patients' waiting time for surgery. A low acceptance rate by SNS patients seems to be the root cause of this public health policy's overall unsatisfactory impact. In this chapter, the methods employed in order to fully dissect this problem will be discussed.

As an initial goal, this research intended to model which patients' individual aspects determined their decision to either accept or decline an SV, and to which extent that influence was significant. The HBM model was, therefore, adapted to translate this specific health-related context (Subchapter 4.1). Data regarding patients' personal perceptions and beliefs concerning SVs were collected by a two-part questionnaire – a "measurement instrument" created to indirectly assess the conceptual model's variables among a study sample, in representation of the target population. The validity of such model to characterize the SVs' acceptance problem was evaluated through Structural Equation Modelling (SEM) and its encompassing multivariate statistical methods (Subchapter 4.2).

The following steps of this research were based on exploring a simulation model of the Portuguese surgery waiting list, developed with AnyLogic software (Subchapter 4.3). Such tool enabled the performance of multiple impact analyses resultant of different behavioural changes, which could enlighten policymakers regarding effective ways of improving SVs' implementation results.

4.1 Conceptual model and research hypotheses

A conceptual model was proposed to represent and explain patients' individual decision to accept or decline an SV while enrolled on the surgery waiting list. From the three health-related psychological models explored in Chapter 3, the HBM model demonstrated to be an adequate choice, since it was originally created to be applied in the healthcare field and, therefore, seemed to include all important psychological and physical aspects involved in one's internal decision-making process regarding personal health matters, like the one of interest in this study.

All constructs of the model were included: determinant dimensions – perceived susceptibility, perceived severity, perceived benefits, perceived barriers, cues to action, self-efficacy, and health motivation –, as well as the health behaviour itself – in this case, measured in terms of intention to behave as a predictor of actual behaviour. Since the behaviour being studied was SVs' acceptance, HBM's constructs were applied to that context. Threat dimension (perceived susceptibility and severity) was considered to be related to people's thoughts regarding surgery waiting times beyond the established TMRGs and their possible impact on general health condition and quality of life. Perceived benefits referred to the advantages accepting an SV could bring to patients' lives, while perceived barriers included the main reasons for SVs' refusals reported by SNS patients in a SIGIC's enquiry study [19]. Cues to action comprised third persons' opinions and inputs regarding the subject (family, friends, or family practitioners). Self-efficacy portrayed patients' capacity to communicate their intention to accept and to actually make use out of an SV. Health motivation was related to someone's overall interest regarding personal health. Finally, intention to behave represented patients' willingness to accept an SV, if offered one in the future.

In general, if patients think of themselves as more susceptible to a possible health deterioration following a long waiting period, if they consider that such a long wait can lead to severe disease, if they value being treated faster and believe an SV would make that happen, if they are in an environment that incentivizes SV's acceptance, if they think they are capable of performing the actions involved in accepting an SV, and if they are highly concerned with their personal health, then it is expected that they would want to accept an SV if offered one. On the other hand, if patients consider that such decision involves many obstacles and that it has a lot of disadvantages, it is expected that they would not accept the transfer offer. Following that reasoning, seven research hypotheses were defined regarding the nature of the causal relationships between the variables of the model and intention to behave, as follows:

- H1: Perceived susceptibility is positively associated to intention to accept an SV;
- H2: Perceived severity is positively associated to intention to accept an SV;
- H3: Perceived benefits are positively associated to intention to accept an SV;
- H4: Perceived barriers are negatively associated to intention to accept an SV;
- H5: Cues to action are positively associated to intention to accept an SV;
- H6: Self-efficacy is positively associated to intention to accept an SV;
- H7: Health motivation is positively associated to intention to accept an SV.

The diagram depicted in Figure 7 represents the HBM conceptual model adapted to the SVs' acceptance problem, as well as the above-stated research hypotheses.

Along the next subchapter, the methodology employed during the quantitative study carried out in order to examine and validate hypotheses H1 to H7 is going to be explained.

4.2 Quantitative study

A quantitative study was conducted to validate the model of Figure 7 in the presented context. A newly created HBM-inspired questionnaire enabled an indirect measurement of models' constructs among a sample of Portuguese citizens. The model was continuously adapted to fit the collected data and variables which did not satisfactorily reflect the problem of study were discarded, so that reliable causal estimates of the model could be obtained and the main determinants of SVs' acceptance revealed.

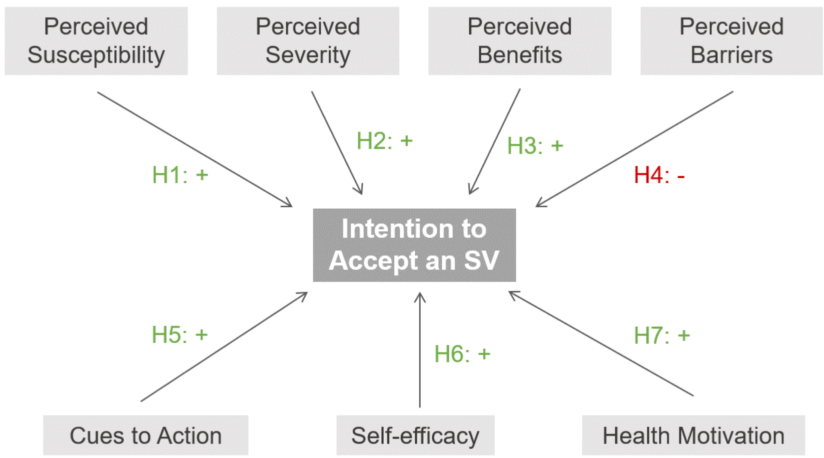

Figure 7: Conceptual HBM model for SVs' acceptance behaviour and its corresponding research hypotheses. Hypotheses written in green and marked with a plus sign denote positive associations between the model's constructs, while H4 is written in red and marked with a minus sign to reinforce its negative nature.

4.2.1 Gathering of participants

In an ideal scenario, the present study's questionnaire would have had as its target population a set of patients currently enrolled on the Portuguese surgery waiting list. That would have resulted in a more realistic sample representation and allowed respondents to answer accordingly to their real perceptions regarding the matter at hand, since they would be actually going through the waiting process and would be eligible to receive an SV.

Faced with the impossibility of directly approaching the national waiting list's population, due to privacy legislation, the questionnaire developed under the scope of this dissertation was distributed among the general population. Despite being a slight adjustment, it should not significantly compromise the results as, in the end, everyone is susceptible of entering the waiting list for surgery and be presented with the option to use an SV. Nevertheless, this means that people who had never found themselves in a similar situation had to imagine that particular scenario and try to answer in harmony with what they presently think their perceptions on the subject would be. People who already had the experience of being offered an SV in the past should have been able to answer accordingly to what they recall from that time. Due to the same limitation, behaviour per se could not be measured and it had to be replaced by intention to behave, which has already been done in other HBM studies [48, 49].

The survey was developed on an online format, using Google Forms, and shared on different social networks (Facebook, LinkedIn), as well as distributed door-to-door. Even though an online spread strategy enables researchers to easily engage with a lot of different people, promoting a high number of responses, it was felt the need to complement this approach with a door-to-door enquiry process in order to reach the elderly, that, despite not having a significant online presence, represent a very important part of this study's target population.

Answers were collected from June 28th 2021 to August 25th 2021. A pre-test was performed by presenting to a few people an initial version of the survey. That allowed to validate questionnaire items, their corresponding measurement scales, and provided valuable information regarding important adjustments to be made for the final version of the survey (that will be explicitly justified in the next subchapter).

4.2.2 Questionnaire structuring

The questionnaire here developed, which can be found in Appendix A, consisted of two main sections: an initial part, that entailed some brief demographic questions to better characterize the sample being studied (age, gender, district, education level) as well as three other items concerning subjects' overall health perception and past experience on the surgery waiting list, and a more extensive second part, which was intended to capture and measure HBM's main constructs. All questions were written in Portuguese.

Each construct of the model, that from now on will also be designated by latent variable, represents an unmeasurable dimension which, therefore, needs to be indirectly assessed through subgroups of observable variables with which it maintains some kind of causal relationship [50–52]. These measurable variables (also called indicators) constitute the survey's items, from whose answers it is possible to infer information regarding the construct they relate to. Therefore, each HBM construct is represented in the questionnaire by a set of questions with which it shares some association – the measurement items' choice should guarantee that different aspects of each construct are covered, so that when those related measures are combined it becomes possible to fully "operationalize" the latent variable they characterise [50, 53].

Thus, in this last part of the questionnaire, respondents were faced with 26 enquiry items, organized into seven groups, corresponding to the main constructs of the model in place: perceived susceptibility (measured by two items of the survey), perceived severity (measured by four items), perceived benefits (three items), perceived barriers (six items), cues to action (five items), self-efficacy (two items), and health motivation (four items). To answer, respondents had to select their level of agreement with each item statement on a five-point Likert scale (ranging from completely disagree, level one, to completely agree, level five). Table 4 presents the questionnaire items that composed this second part of the survey, along with their identification number, measurement scale, and the HBM construct they referred to. As it may be observed, each construct was determined by more than one observed variable, as it is recommended in the literature. It has been shown that having at least three indicators per latent variable leads to more statistically significant results during the analysis of any measurement instrument [53].

The majority of the questions was adapted to this project's context from the literature, where similar surveys were used to develop HBM models to other health-related issues, such as vaccine intake [48, 49], tumour screening [40], smoking [54] or oral hygiene habits [39]. Particularities of the SV's system, namely SNS patients' main reasons for SV's refusals [19], that were already explored in Chapter 2, were also very important to create measurement variables that truly fit the situation under study.

In addition, a final question was used to enquire participants regarding their willingness to accept an SV if presented with that hypothesis ("If you would receive a surgery voucher in the future, would you be willing to use it?"), in order to measure intention to behave, the ultimate dimension of the HBM model. Initially, during the pre-test stage, this item was implemented with a dichotomic response scale, so that respondents could choose between “yes” or “no”. At the time, this was thought to be a suitable approach since it was supposed to capture individuals’ intention to engage in a dichotomic behaviour (one either accepts or does not accept an SV, by declining it or simply not answering it), and based on the fact that a similar method had been used in other HBM studies found in the literature, concerning vaccination intake, for example. However, this type of binary answer proved to not be very useful once an analysis of these preliminary results was performed, since 98% of the enquired sample at that stage had selected the option “yes”. That meant that no significant causal relationships could possibly be found between any of the HBM constructs and the final intention to accept an SV.

Furthermore, in cases where the survey was distributed personally door-to-door, it was possible to collect some feedback regarding that same question: although the majority of people tended to answer “yes” when confronted with only two options (“yes”/”no”), they showed varying confidence levels in their answer (meaning that people were, in general, prone to accept the SV, but assuming that some conditions they individually value would be met in that circumstance). Therefore, to try to capture those different levels of willingness to accept, the originally two-option final question was adapted to a five-point scale, where category one referred to “definitely no” and category five to “definitely yes”.

Nevertheless, the results obtained from the initial version of this questionnaire item did not entirely go to waste. The overwhelming value of 98% of willingness to accept, when compared to the 18.8% SV’s acceptance rate registered in 2019 (see Table 3), shows a tremendous potential for this system to achieve significantly better results in the future, if adjustments are made and the conditions offered to patients actually meet their personal expectations.

4.2.3 Data analysis

Once a questionnaire is finalized and the answers are all collected, it is time to validate and interpret the obtained results. When it comes to validity, it is important to emphasize that it does not concern the measurement instrument by itself, but rather if theory and evidence align and support its use on a specific context – in other words, if by using a certain instrument researchers are truly measuring what they intended to assess [50]. While theoretical evidence based on similar studies found in the literature [37, 48, 49, 55] seemed to support the validity of the survey here employed and its corresponding model, it was crucial to search for other sources of validity proof.

That is where SEM came into play: “a general modelling technique, used for testing the validity of theoretical models defining causal hypothetical relationships among variables” [56]. SEM encompasses a number of different multivariate statistical methods – such as variance and covariance analyses, multiple regression, factor analysis, path analysis, etc. –, which may be combined and used to ultimately determine the representative parameters of the causal influences existing between independent and dependent variables that constitute the model [56–58]. With SEM, such correlations among multiple variables can be estimated simultaneously. Besides that, it enables researchers to consider both observable and unobservable (latent) variables, the latter being indirectly operationalized by the former – a very important tool when dealing with abstract and theoretical concepts (such as behaviours, emotions

Table 4: HBM constructs of the model, along with their questionnaire items' identifiers, their corresponding statements and measurement scales.

HBM Construct	Questionnaire Item	Statement	Scale
Perceived Susceptibility	Q1	It is likely that my health status significantly deteriorates if my surgery is not performed within the TMRG provided by law.	From 1 (completely disagree) to 5 (completely agree)
	Q2	A long waiting time may significantly compromise my post-operative outcomes and my recovery.	
Perceived Severity	Q3	Health complications resulting from a long waiting time can severely affect my ability to work.	
	Q4	Health complications resulting from a long waiting time can severely affect my ability to perform common daily tasks.	
	Q5	Health complications resulting from a long waiting time can severely affect my family life.	
	Q6	Health complications resulting from a long waiting time can severely affect my quality of life.	
Perceived Benefits	Q7	Accepting the surgery voucher means that my health problem will be solved faster.	
	Q8	Accepting the surgery voucher will enable me to regain my previous quality of life more quickly.	
	Q9	Accepting the surgery voucher will decrease the likelihood of deterioration of my health stratus.	
Perceived Barriers	Q10	I would not like to be treated by a medical team other than the one from my home hospital.	
	Q11	I would not like to be treated in a hospital other than my home hospital.	
	Q12	Going to a hospital outside my residence area is too much of an inconvenience for me.	
	Q13	Having surgery in a hospital outside my residence area is too much of an inconvenience for me.	
	Q14	My work and/or family situation would cause me to postpone my surgery as much as possible.	
	Q15	The anxiety resulting from the surgery date's approaching would cause me to postpone the procedure as much as possible.	
Cues to Action	Q16	Family members' support to accept the surgery voucher would influence my decision.	
	Q17	Friends' support to accept the surgery voucher would influence my decision.	
	Q18	The support of my home hospital's medical team or my family doctor to accept the voucher would influence my decision.	
	Q19	Contacting with other patients who had previously scheduled their surgeries though the surgery vouchers' system would influence my decision.	
	Q20	A deterioration of my health condition during the waiting period would influence my decision.	
Self-efficacy	Q21	Upon receiving a surgery voucher, I would be able to contact and schedule the surgery myself at one of the hospitals from the options list.	
	Q22	If it was not clear to me how I should proceed to use the voucher, I would be able to search information regarding that process.	
Health Motivation	Q23	I try to maintain a healthy lifestyle.	
	Q24	I try to maintain a regular physical activity.	
	Q25	I try to maintain a healthy diet.	
	Q26	I regularly search for information to help me stay healthy.	
Intention to Accept	Q27	If you would receive a surgery voucher in the future, would you be willing to use it?	From 1 (definitely no) to 5 (definitely yes)

or attitudes), which justifies why SEM became such an appreciated analytic method in many research areas, especially surrounding social sciences [56–58].

Factor analysis, a vital part of SEM, can be used to empirically explore the internal dynamics and relationships between a set of observed variables (surveys' items). Thus, it is possible to analyse if/how certain subgroups of measures might be correlated and, when combined, responsible for characterising a common factor (or latent variable) [50]. There are two main procedures that can be used, individually or combined, to perform this kind of evaluation: Exploratory Factor Analysis (EFA), which explores an underlying structure present in some dataset without any pre-imposed form, and Confirmatory Factor Analysis (CFA), which confirms the adequacy of a data sample to a predefined model.

Although it is possible to find applications of similar enquiry instruments to somewhat related problems in the literature, it is, at this dissertation's author best knowledge, the first time the HBM model is used to explore SV's general acceptance, as well as the first time this specific questionnaire is employed in a study. As a newly developed measurement instrument being applied to a novel context, the literature suggests that a full factor analysis must be performed: starting with an EFA to understand the internal structure of survey's responses, followed by a CFA in order to corroborate it [50].

On the other hand, path analysis can be used to study the intensity of causal relationships among latent variables, based on an already established structure. In fact, what this method does is to attribute quantitative meaning to predefined (based on theoretical evidence and previous knowledge) causal interactions, rather than uncover such causes per se [56].

It is important to mention that, although being a type of factor analysis, EFA is usually not considered a part of SEM, due to its exploratory nature. According to Marôco [56], SEM is a framework of statistical techniques driven by theory: researchers should collect data and check if it confirms or not a predefined theoretical model. On the contrary, EFA is data-driven, as it searches for a theoretical proposition that fits the gathered information [56]. Nevertheless, both kinds of analyses are useful at different stages of a research study and, therefore, should be joined whenever necessary [50].

Usually, SEM's general model encompasses two sub-models: a measurement model, that is used during CFA and establishes how the multiple sets of observed variables translate the latent structure present in the overall model, and a structural model, which defines latent variables' causal relationships and is a result of path analysis.

Along the quantitative study conducted in this dissertation, both data-driven and theory-driven techniques were employed. The study started with an EFA to explore the embedded structure in the measurement instrument's answers. The factor model here revealed was then used during SEM – first, a CFA was performed in order to confirm the EFA-obtained structure, and then path analysis was used to determine the causal relationships among the variables that composed the HBM-inspired model.

Sample and data considerations:

Answers from residents in Portugal's autonomous regions of Azores and Madeira were excluded from the studied sample, since these territories have autonomous regional health services and manage their own surgery waiting lists independently from the rest of the country.

Regarding the suitable sample size to perform factor analysis, there is a lot of discussion in the literature. While earlier recommendations indicated a minimum number of participants such as 100 or 200, or even a ratio of five or ten participants per parameter to estimate, more recent studies suggest that sample dimensions are highly dependent on the unique characteristics of measurement instruments and models (number of factors, number of variables per factor, strength of items relationships with their corresponding factors, etc.).

Mundfrom and Shaw [59] developed guidance tables for minimum necessary sample sizes, depending on the number of factors and level of communalities (amount of variables' variance accounted for a certain factor) present in a factor analysis. For the conditions experienced in the present research, the authors suggested a minimum of 160 to 170 participants, which is satisfied in this study. Nevertheless, factor analysis is considered a large-sample procedure, whose results tend to benefit from larger sample pools [53].

Best practice guidelines for conducting factor analysis recommend that exploratory and confirmatory studies should not be performed in the same dataset, as a sequential study of EFA and CFA using the same sample would be of little value – in that case, CFA would most certainly corroborate the former analysis' model results. Therefore, independent samples are required in order to test an instrument's true robustness [50, 56].

A common approach used when EFA and CFA are joint in a same study is cross-validation: a sample is randomly divided into two independent subgroups, one used in EFA and the other in CFA, in order to insure that the obtained results are not sample-specific, that they are replicable and reliable [50, 60]. That was also the strategy followed in this dissertation, where the collected answers from the questionnaire were randomly divided in half (approximately) to form two independent subgroups, one for each factor analysis technique.

Item Q27 was not considered for EFA nor CFA (although it was reinserted further ahead, for path analysis), as it does not represent any latent factor of behaviour, but rather a measure of one's overall willingness to behave (which, in this case, was an approximation used to measure SVs' acceptance behaviour).

It is also important to mention that no missing data was found in the obtained total sample.

Characterization of the sample:

A sociodemographic analysis of the obtained sample was performed in order to assess if it was representative of the actual target population, by comparing it to the sample utilized in a SIGIC's previous study that aimed at understanding SVs' refusal motives. Representativeness is important to allow extrapolation of the here attained results to the whole population of interest (formed by all patients enrolled on the surgery waiting list).

Age, gender, ARS, and educational level were the variables considered in the initial part of the questionnaire. Fisher's exact test of independence was employed to confirm if there was some degree of dependence between any of these sociodemographic characteristics and intention to accept an SV. To simplify this comparative evaluation, answers regarding intention to accept (Q27, the last question of the survey) were, just for this initial stage of the analysis, aggregated from a five-point Likert scale and divided in only two groups: "intending to accept", which included answers from levels three to five, and "not intending to accept", that considered replies of both values one and two.

The choice of Fisher's test was due to its adequacy to considerably small-sized samples, like the one used in this study, and the fact that the variables to be compared were all nominal [61]. Results with p-values above the alpha level of significance (defined at 0.05) implied a rejection of the null hypothesis, meaning that no relevant dependence relationships were found.

The following survey question referred to respondents' self-evaluation of their health status, on a scale from zero (very poor) to ten (excellent). Since this was an ordinal variable, Fisher's test did no longer apply. Therefore, to perform an independence evaluation of this ranking item and willingness to use an SV, the Mann-Whitney U test was used [62].

To conclude this first part of the questionnaire, participants were asked about their previous history on the surgery waiting list and past experience using SVs ("Have you ever been enrolled on the surgery waiting list before?", "If yes, have you ever received an SV? Did you use it?"). As these are, once again, nominal variables, Fishers' exact test of independence was used to evaluate the existence of any relevant dependencies between these items and HBM's intention to behave.

After this initial analysis, and for the rest of the study, answers to the last questionnaire item regarding intention to accept an SV were used back in its original five-point ordinal scale format.

EFA:

Child [63] defines EFA as "a statistical technique used to explore a possible factor structure underlying a set of observed variables, without imposing a preconceived structure on the outcome". When performing this kind of analysis, researchers' main goal is to detect possible correlations among the variables they originally measured, following an exploratory approach, as the name suggests. If a certain subset of observed measures presents high correlations among themselves, that is an evidence of some kind of subjacent structure: that group of variables must be affected by a common latent (unobserved) variable [51, 53]. This common factor influences individual scores on each measurable indicator (survey item) and accounts for the covariance (or shared variance) among them. The non-shared variance between a set of indicators of a same construct is attributed to measurement errors, δ_i. Factors' correlations, ϕ_{jj}, may also be computed during EFA [53].

Besides detecting correlations (which are equivalent to standardized covariances, ranging from -1 to 1), EFA is also useful to identify indicators that do not show an empirical fit to the construct they were initially thought to be related to. These items should, therefore, be deleted from the survey and other additional analyses, unless there is enough theoretical evidence that justifies their maintenance in the study [50].

Since it does not require any predefined structure, this kind of analysis is commonly utilized by researchers, for example in social and behavioural sciences, at the early stages of a study, when developing new models, theories, and applications – new measurement instruments are created, based on fundamentals extracted from literature or interviews, and evaluated through EFA [50, 51, 53].

Concerning this dissertation, the conceptual model developed was not built from scratch, as it is essentially an application of the already established HBM model, largely explored in the literature and on an extensive amount of investigation articles, to a new practical context. Therefore, in this case, the model's constructs were a priori identified and the questionnaire items were either created or adapted to measure SV's acceptance (largely based on surveys developed in previous related research, as already explained). Nevertheless, the fact that this exact questionnaire had never been utilized before and its base model had never truly been applied to this specific context justifies, according to many authors [51, 53], that an EFA analysis is performed in order to understand if the newly developed survey items actually measure the model dimensions they were intended to measure and if those variables do have an impact in the context to which they were here being specifically applied to [51].

During EFA, relationships between observed variables and a set of factors are studied, through the computation of factor loading estimates (variable-factor correlation coefficients), λ_{ij}. Indicators that share high covariance also present significant loading coefficients on their common construct, being, therefore, collapsed into that same factor. If the measurement instrument turns out to be valid, the identified factor structure would meaningfully represent HBM's dimensions. Therefore, EFA can be also defined as a technique that "attempts to determine the smallest number of constructs that parsimoniously explain the covariation observed among a set of measured variables" [53]. Choosing the ideal number of factors to consider in the analysis is just one of many dilemmas researchers need to surpass along the EFA process.

The EFA technique consists of a sequence of methodological steps for which there are more than one approach available for researchers to decide on. Examples of that are the factor estimation method or the type of rotation employed in the analysis, which may vary from study to study. Although there are some guidelines for best practice in the literature, there is still much controversy and debate surrounding many of those recommendations, leaving it to researchers to choose accordingly to what they think is the best strategy to follow depending on their studies' specifications [50, 53, 64]. As this dissertation is no exception, next it will be described the EFA procedure decisions made to best fit the characteristics of this particular project.

The first step in this type of statistical studies must always be to check if the collected dataset is actually appropriate to perform factor analysis, i.e., if there are sufficiently significant intercorrelations among the observed variables. The two most employed tests for this purpose, and that were also used in this study, are the Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) test [53]. While the former evaluates the factorability of the correlation matrix (tests if such matrix is an identity matrix, in which case the variables would not share any type of significant relation and, therefore, a factor structure would not be present), the latter examines sampling adequacy (measuring the extent to which observed variables' correlations are a result of common latent variables). In order to be considered factorable, a sample must produce a significant chi-square value (with a p-value below alpha, 0.05) on Bartlett's test and a KMO estimate higher than 0.50 (according to Kaiser [65], values higher than 0.90 are "marvellous", 0.80 "meritorious", 0.70 "middling", 0.60 "mediocre", 0.50 "miserable", and under that mark considered unacceptable).

Regarding correlation values among variables, although Pearson coefficients are the usual choice, they will not be used in this analysis, since they assume that the measures extracted from the instrument being studied are normally distributed along a continuous scale [56]. On the contrary, the variables considered in this questionnaire were measured on an ordinary scale (Likert scale), as it happens in the vast majority of behavioural studies. For these cases, and especially when such scales are not very large, literature suggests the use of polychoric correlations instead [50, 53]. Polychoric correlations among two ordinal items, x_1 and x_2, are computed based on the interpretation that such variables are an approximation of two subjacent continuous variables, with normal distributions, ξ_1 and ξ_2, respectively. Ordinal variables implemented by a five-point Likert scale, like the ones used in this study, can be conceptualized in the following form:

$$x_i = \begin{cases} 1 - Completely\ disagree,\ if\ \xi_i \leq \xi_{i1} \\ 2 - Disagree,\ if\ \xi_{i1} < \xi_i \leq \xi_{i2} \\ 3 - Neither\ agree\ nor\ disagree,\ if\ \xi_{i2} < \xi_i \leq \xi_{i3} \\ 4 - Agree,\ if\ \xi_{i3} < \xi_i \leq \xi_{i4} \\ 5 - Completely\ agree,\ if\ \xi_i > \xi_{i4} \end{cases}, \tag{1}$$

with $i = 1, 2$ and where the five ordinal categories are approximated to intervals of the continuous variables, defined by thresholds ξ_{i1} to ξ_{i4} (by convention, $\xi_{i0} = -\infty$ and $\xi_{i5} = +\infty$). The polychoric coefficient between the two latent continuous variables is computed in representation of the correlation among the ordinal variables, as Figure 8 illustrates [56].

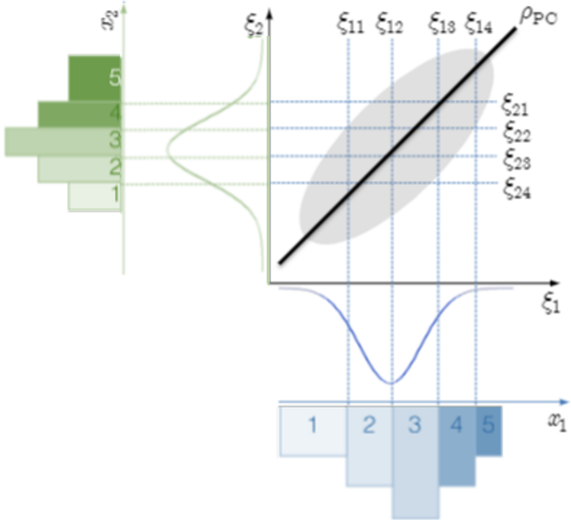

Figure 8: Illustration of the polychoric correlation, ρ_{PC}, between two latent continuous variables, ξ_1 and ξ_2, operationalized by two observed five-point ordinal variables, x_1 and x_2. ξ_{i1}, ξ_{i2}, ξ_{i3}, and ξ_{i4} denote the continuous intervals' thresholds that correspond to the five categories of the categorical scale [56].

Furthermore, it has been shown that, when working with real datasets, violations of normality are quite normal. Skew (referring to the symmetry of the measures of a variable) and kurtosis (associated to the relation between the height and width of measures' distribution) values can be used to evaluate deviations from normality. Skewness higher or equal to $|2.0|$ as well as kurtosis higher or equal to $|7.0|$ are strong indicators that normality should be discarded [50, 53]. In addition to the analysis of skew and kurtosis values, Mardia's estimates, a common test used to assess multivariate normality, were also used to search for signs of a non-normal distribution among the surveys' dataset [50, 53].

The chosen method to estimate the common factors present in the data was principal axis factoring. Once again, due to the non-normal and ordinal nature of the observed variables, this is the most recommend method in EFA's best practice guides [50]. It is an iterative approach based on least-squares estimation that does not make any distribution assumptions on the data. The result is a list of possible factors to be extracted from the analysis. Each factor is accompanied by its corresponding eigenvalue, which is an estimate of the amount of data variance it can explain on its own (independently from the other factors). Factors are presented in descending order of their eigenvalues, therefore assuring that the first factor is responsible for the highest amount of variance, and so on [53].

Different criteria can be used to decide on the ideal number of factors to maintain in the model. A compromise value should be found, so that the number of factors considered is big enough to capture the model's fundamental structure (covering a significant part of the total variance) while small enough to not represent a computational and interpretational burden [66]. Methods like Kaiser's "eigenvalue-greater-than-one" rule (where all factors with an eigenvalue higher or equal to one are selected) and the visual scree (a graphical approach that implies all factors displayed before a convex point in a plot of their eigenvalues' ordered magnitude should be considered) were commonly used in the past; however, they are now considered a bit outdated [53, 66]. On the other hand, parallel analysis (the selected approach to be applied in this study) has emerged as a reliable method for factor extraction. It is based on a random generation of an eigenvalues list that is compared to the one obtained for the data sample. Factors whose experimental eigenvalue is higher than its corresponding random value are considered adequate to be taken into account in further analysis [53, 66]. In this study's EFA, despite mainly using parallel analysis, eigenvalues' scree plot was also examined as a way of corroborating the number of factors estimated by the former method and enhancing the confidence in the obtained value.

As sometimes the number of factors to include in this analysis is not clear, it is recommended that, in case of doubt, researchers start with the highest plausible number, evaluate the proposed solution, and repeat the same process decreasing the number of factors one at a time. The best solution should determine the final number of factors to consider [53].

Once the ideal number of extracted factors is determined, the EFA should be continued considering that decision. In order to facilitate output interpretation, it is usual to perform factor rotation. Such procedure improves theoretical meaning of the results, where correlations of the variables on each factor are estimated [53]. There are essentially two streams of rotation methods: orthogonal or oblique. The main distinction between the two is that orthogonal rotation tends to be applied when there is strong evidence supporting the inexistence of correlations among the factors considered, while oblique rotation is adopted when such information is not obtainable or when researchers have reasons to consider that factors may indeed have some kind of correlation [50, 53]. Given the nature of the problem being analysed in this dissertation and the fact that behavioural dimensions are plausible to interrelate, there was a possibility that correlations between some latent variables did exist. Therefore, it seemed appropriate to select oblique factor rotation as the way to go. Within this type of rotation, different mathematical methods can be followed – for example, *promax* and *oblimin* –, although they all appear to lead to similar results [50]. For this specific study, *oblimin* was the chosen oblique rotation technique.

Through oblique rotation, two kinds of factor loadings could be obtained: structure and pattern coefficients. The former represent correlations between common factors and their corresponding measured variables, while the latter characterize similar relationships this time when the effects of all other factors in such variables are discarded [50]. Therefore, structure and pattern coefficients should be similar in case factor intercorrelations are not significant. During this analysis, only pattern coefficients were evaluated, since the point here was to define items as dependent on only one factor.

The cut-off value that determines which pattern coefficients are considered significant loadings of the variables among a set of considered factors is also not consensual in the literature. While Hair [67] stated that variables should attain pattern coefficients higher than 0.40 on a certain factor in order to be considered significant for the analysis, some more recent studies have been considering more restrictive values, up until 0.70 [53]. In this dissertation, 0.50 was the defined minimum loading limit for pattern coefficients to be considered significant. According to Brown [68], during an EFA analysis, researchers should aim at finding a solution where each factor is significantly loaded by several variables, which, in turn, present significant loading values on only one factor (and negligible values on all others). Therefore, variables should be excluded from further analysis when they do not present significant loadings on any factor (i.e., if factor loading is under 0.50) or when they present significant loadings on more than one factor (cross-loadings). Despite all that, in some cases theoretical arguments may still be invoked to justify the inclusion of such "problematic" variables in the study [50]. It is also important to mention that, after the decision to remove certain items is made, EFA should be redone without those variables to ensure model's validity is not lost.

Besides factor loadings, common factor analysis also produces separate variances for each variable: a communality term, h_i^2, and a uniqueness term, u_i^2. Communality considers the part of a variable's variance which is accounted for by every factor considered in the EFA model (and not just by the factor where it is heavily loaded), while uniqueness refers to the part of an item's variance that cannot be explained by any part of the model (it is usually referenced as an error component although it emerges not only from measurement errors, but also as a result of existent nonshared variance). Low communalities, typically under 0.40, may also be used as an argument for items' elimination from a study, since it indicates a poor representation of those variables by the developed model [50].

Figure 9 shows a conceptual representation of a simplified EFA model, with only six indicators and two factors. The model can also be written as:

$$\mathbf{x} = \mathbf{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}, \tag{2}$$

in this case, also equivalent to:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} & \lambda_{32} \\ \lambda_{41} & \lambda_{42} \\ \lambda_{51} & \lambda_{52} \\ \lambda_{61} & \lambda_{62} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{bmatrix}, \tag{3}$$

where $\mathbf{x}$ is the independent observed variables' vector, $\mathbf{\Lambda}_x$ is the loading coefficients' matrix of variables $\mathbf{x}$ in factors $\boldsymbol{\xi}$, $\boldsymbol{\xi}$ is the independent latent variables' vector (or factors' vector), and, lastly, $\boldsymbol{\delta}$ is the measurement errors' vector of variables $\mathbf{x}$ [56]. The correlation matrix of factors ξ_1 and ξ_2 is given by:

$$\mathbf{\Phi} = \begin{bmatrix} 1 & \\ \phi_{21} & 1 \end{bmatrix}. \tag{4}$$

For an EFA solution to be considered acceptable, there is still one more criterion to be met. Each factor should obtain a Cronbach's alpha coefficient higher than 0.70, in order to prove to have internal consistency [56]. This measure examines if the whole subset of items concerning a certain latent variable, or factor, measures the same thing consistently, based on their average correlations. However, this index does not ensure researchers are truly measuring the construct they intended to – Cronbach's alpha indicates reliability, not validity. Nevertheless, both concepts are intertwined, since high reliability is a non-sufficient but required requisite for validity [50]. According to George and Mallery [69], alpha

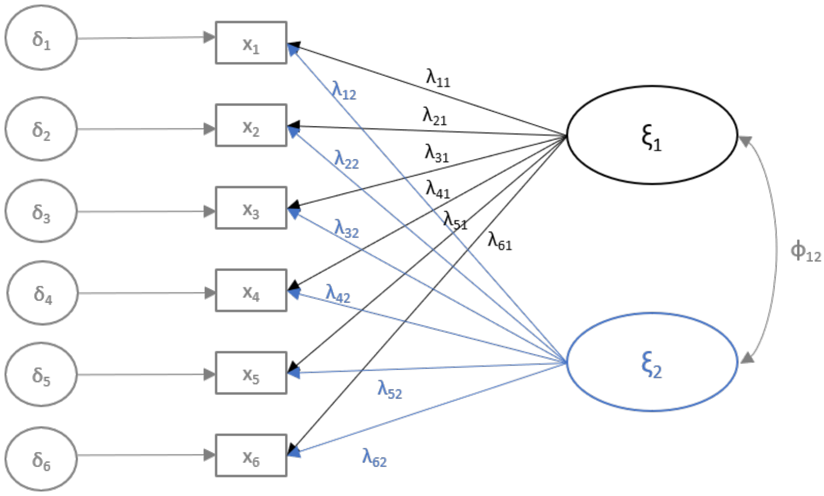

Figure 9: Conceptual representation of an EFA considering six observed variables (illustrated by rectangular shapes), x_1 to x_6, and two factors (represented by ovals), ξ_1 and ξ_2. Arrows represent construct-indicator causal relationships and λ_{ij} their corresponding factor loadings. Double-ended arrows stand for correlations between factors, ϕ_{jj}. Error components, δ_i, are also considered. (Adapted from [50].)

coefficients higher than 0.70 are acceptable, higher than 0.8 are good, and above 0.9 are considered excellent – which was the reference scale considered in this study.

CFA:

Unlike EFA, CFA is performed when a structure between the observed variables has been previously identified and explored. Therefore, it is possible to use CFA to confirm pre-established correlations between observed and latent variables, based on the factor structure uncovered during EFA [50, 51]. Thus, in order to carry out a CFA, certain characteristics of the model have to be prespecified, such as the number of factors to consider and which items load on each factor [51, 56].

As it is mainly used to evaluate how a certain model adjusts to empirically observed correlations between sets of items, CFA often consists on the initial step of a SEM analysis [56]. The model used during a CFA is usually designated by measurement model and it specifies each factors' set of corresponding indicators (x_i), as well as their loading coefficients (λ_{ij}). Besides that, this model also includes error components (δ_i), and correlations among different factors (ϕ_{jj}) [56, 58]. Error components are also considered latent variables, since they cannot be directly measured, and account for external causes influencing an observed variable that not the factor it is related to in the model [56]. Therefore, such errors can be defined as:

$$\delta_i = 1 - R_i^2, \tag{5}$$

where R_i^2 stands for an item's determination coefficient (the proportion of that item's variance explained by its respective factor in the model). Some authors argue that R_i^2 is a measure of items' internal reliability (ideally close to one), while others attribute it to the square value of an indicator's standardized loading factor, λ_{ij}^2. Nevertheless, in general, R_i^2 and λ_{ij}^2 assume approximate values for a same item *i* [56, 58].

For a CFA model to be operationalizable, it must be overidentified, i.e., it must present a higher number of variances observed in the data than parameters to be computed. Such condition can be assessed through the model's number of degrees of freedom, that must be positive [50].

Figure 10 depicts CFA's conceptual representation, following the notation previously applied to EFA. This model is also defined by equation (2), with the difference that, this time, each observed variable is loaded on only one factor, meaning that its loading coefficients' matrix is given by the following structure:

$$\mathbf{\Lambda}_x = \begin{bmatrix} \lambda_{11} & 0 \\ \lambda_{21} & 0 \\ \lambda_{31} & 0 \\ 0 & \lambda_{42} \\ 0 & \lambda_{52} \\ 0 & \lambda_{62} \end{bmatrix}. \tag{6}$$

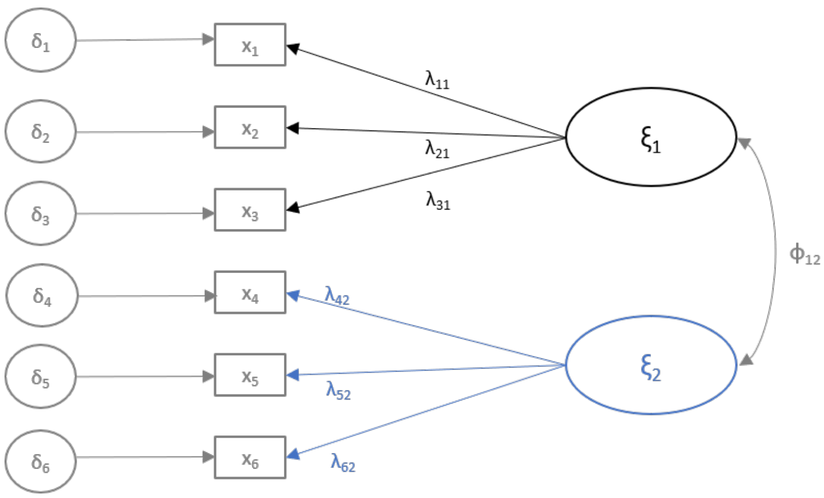

Figure 10: Conceptual representation of a CFA considering six observed variables (illustrated by rectangular shapes), x_1 to x_6, and two factors (represented by ovals), ξ_1 and ξ_2. Arrows represent construct-indicator causal relationships and λ_{ij} their corresponding factor loadings. Double-ended arrows stand for correlations between factors, ϕ_{jj}. Error components, δ_i, are also considered. (Adapted from [56].)

During a CFA, variances can be extracted from the data through multiple mathematical methods, just like it happened in EFA. In this case, where the measurement scales were ordinal and a non-normal distribution of the data was plausible (that was, nevertheless, examined by skew and kurtosis indexes), the weighted least squares' estimators family is the most suggested in the literature, since it does not assume any kind of data distribution and has the capacity to compute polychoric correlations [50]. Diagonally weighted least squares was the specific method chosen for this analysis, also due to its capacity to obtain satisfactory results even when working with considerably small samples [58, 70, 71].

Multiple model fit indices can be used to assess how well a sample supports an hypothesized conceptual model. As it is recommended in the literature, for this analysis there were produced different types of fit indices: absolute – χ^2/df –, that evaluate model's adjustment per se and without establishing comparisons; relative indices – Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) –, which compare the current adjustment with the best (all observed variables are correlated) or worst adjustment scenarios (all variables are independent); and, finally, an index of populational discrepancy – Root

Mean Square Error Approximation (RMSEA) –, which accounts for the error of approximation to a study's population [50, 56, 72].

The χ^2/df estimate classifies a data-model adjustment as poor when a value higher than five is obtained, acceptable when it is lower than five and higher than two, good if it is between one and two, and very good if it is close to one. Regarding CFI and TLI, values below 0.8 refer to a bad adjustment, between 0.8 and 0.9 it is considered acceptable, between 0.9 and 0.95 good, and higher or equal to 0.95 is very good. Lastly, RMSEA must produce a significant result (p-value below 0.05) lower than 0.10 in order to the adjustment to be considered acceptable – values between 0.05 and 0.10 indicate a good fit, and under 0.05 a very good fit [56].

Factors' composite reliability (CR) was computed to judge the internal reliability of each construct. This estimate also evaluates the consistency with which a set of indicators represents a certain latent variable. Values higher or equal to 0.7 are considered adequate, with one being the maximum reliability level. CR of a *k*-item factor is defined as:

$$\widehat{CR_j} = \frac{\left(\sum_{i=1}^{k} \lambda_{ij}\right)^2}{\left(\sum_{i=1}^{k} \lambda_{ij}\right)^2 + \sum_{i=1}^{k} \delta_{ij}} \ , \tag{7}$$

where each items' error component, δ_{ij}, is given by equation (5) [56].

Regarding the evaluation of this study's validity, to examine if the measurement instrument operationalizes or not SVs' acceptance behaviour, different aspects need to be taken into account: factor validity, convergent validity, and, lastly, discriminant validity [56]. Factor validity respects to items' factor loadings, which should be higher or equal to 0.5 (the same criterion used during EFA). Items' internal validity contributes to factor validity and it's given by the portion of an items' variance explained by its corresponding factor, which should be higher or equal to 0.25. Convergent validity ensures that the majority of items' variance is explained by the factor they are related to in the model. To assess that, factors' average variance extracted (AVE) must be determined, through the following mathematical expression:

$$\widehat{AVE_j} = \frac{\sum_{i=1}^{k} \lambda_{ij}^2}{\sum_{i=1}^{k} \lambda_{ij}^2 + \sum_{i=1}^{k} \delta_{ij}} \ , \tag{8}$$

considering, once again, a construct defined by *k* indicators [56]. Obtained AVE values higher or equal to 0.5 are a sign that this type of validity is verified. To conclude, a model is said to possess discriminant validity when items related to a certain factor are not correlated to any other factors. Such property must be verified by confirming that a factor's AVE is higher or equal to the square values of its correlations with all other factors of the model [56, 58]. During this CFA, all forms of validity were examined.

Path Analysis:

After CFA, where the way latent variables of a model are operationalized by its indicators is defined through a measurement model, SEM proceeds with path analysis, where causal relationships between latent variables are established in a structural model. Upon this analysis, the HBM relational model theoretically developed in Subchapter 4.1 is further explored and, hopefully, validated to explain the behaviour being studied [56].

This last step of SEM simultaneously estimates regression weights for the path trajectories between all constructs of the HBM model and enables a statistical significance analysis of such relationships, specifically between the HBM causal constructs defined during factor structure (perceived threat, perceived benefits, perceived barriers, cues to action, and health motivation) and the ultimate behavioural dimension – intention to accept an SV –, associated to its one measurement item (Q27). Therefore, statistical significance of causal relationships can be evaluated, as well as their impact on behaviour – constructs which have higher regression weights on behaviour are considered to have a bigger influence on patients' intention to accept an SV and, consequently, their final decision to accept it or not.

Below, in Figure 11, it is represented a simple version of the conceptual structural model obtained during a path analysis, considering two independent and one dependent latent variable. Structural coefficients, γ_j, represent the regression weights between causes (exogenous latent variables, ξ_j) and effects (endogenous latent variable, η). This model is defined by the following equation,

$$\boldsymbol{\eta} = \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}, \tag{9}$$

where $\boldsymbol{\eta}$ is the endogenous latent variables' vector, $\boldsymbol{\Gamma}$ is the regression weights' matrix of the causal relationships between exogenous and endogenous latent variables, $\boldsymbol{\xi}$ is the exogenous latent variables' vector, and, lastly, $\boldsymbol{\zeta}$ is the vector of structural models' errors (or disturbances) in relation to the exogenous variables [56].

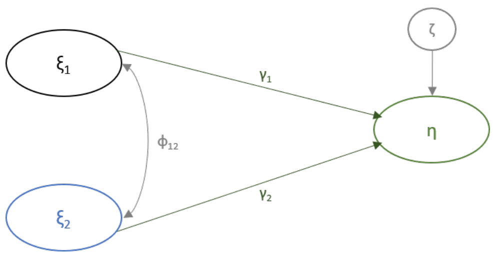

Figure 11: Example of a SEM's structural model, after path analysis, where it is represented the path coefficients, γ_1 and γ_2, between two independent variables, ξ_1 and ξ_2, and one dependent variable, η. Once again, ϕ_{12} stands for the correlation between causal factors, and ζ for an error component (part of η variance that cannot be explained by the independent latent variables present in the model). (Adapted from [56].)

SEM's global model results from the union of both the measurement and structural models, as it is shown in Figure 12. In addition to equations (2) and (9), this model is also defined by:

$$\mathbf{y} = \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\varepsilon}, \tag{10}$$

an equation similar to (2), but corresponding to the measurement model of the dependent observed variables, where $\mathbf{y}$ is the dependent observed variables' vector, $\boldsymbol{\Lambda}_y$ is the loading coefficients' matrix of variables $\mathbf{y}$ in factors $\boldsymbol{\eta}$, $\boldsymbol{\eta}$ is the dependent latent variables' vector (or factors' vector), and, lastly, $\boldsymbol{\varepsilon}$ is the measurement errors' vector of variables $\mathbf{y}$ [56].

Just like in CFA, goodness-of-fit tests (χ^2/df, CFI, TLI, and RMSEA) were performed to the global

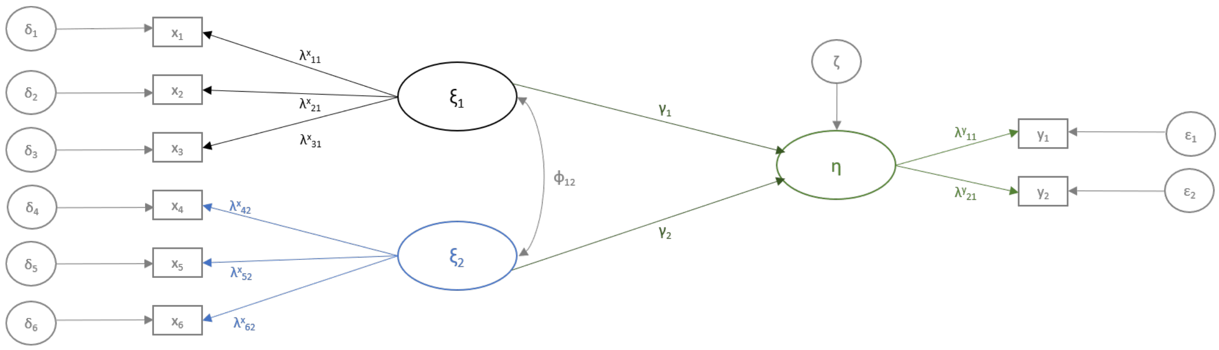

Figure 12: Example of a SEM's global model. This scheme is the result of the union of two measurement models – one of the independent latent variables (Figure 10) and an equivalent one for the dependent latent variable – and the structural model (Figure 11) that represents the relationships between these two types of endogenous variables. (Adapted from [56].)

model obtained after path analysis in order to evaluate the quality of model's adjustment to the data.

Software:

The sociodemographic characterization of the sample, as well as its corresponding independence tests (Fisher's test and Mann-Whitney U test) were performed in R (version 4.1.1). The split of the total sample in two equal-sized subsamples was done using SPSS Statistics (version 26), and later imported to R.

R was also used to perform EFA and both steps of SEM (CFA and path analysis), since it allowed the use of polychoric correlation coefficients, which were the most suitable choice for this study according to the literature. The *psych* package, usually used in social sciences' studies, was an essential tool in order to run the intended EFA. *Lavaan* was the main package used during SEM.

4.3 Simulation study

Within the ambit of the second part of this dissertation's research, an agent-based simulation model of the waiting process for surgery implemented in Portugal was built in AnyLogic software (version 8.7.5). The goal of this simulation was to better understand the effects of behavioural changes in SVs' acceptance levels and in the reduction of the national waiting list. AnyLogic enables researchers to use any combination of the three most known simulation methods – DES, ABS, and SD –, which may be represented through visual modelling languages (e.g., statecharts, flowcharts, stack and flow diagrams, etc.) and complemented with Java code. Portuguese patients on the waiting list for surgery were represented in the model by individual agents, who assumed different states and underwent multiple transitions, representing the dynamics of the waiting process [73].

The model was based on data referring to 2018 and 2019, since these were the last years before elective surgery supply having been affected by the COVID-19 pandemic. During 2019 – the year mimicked during the simulation run –, a total of 724 234 patients entered the surgery waiting list and joined the 244501 that had transited from 2018, according to Table 3. Since these different groups of patients were, at the beginning of 2019, in different points of the waiting process, they had to be modelled according to the different conditions they were in. Therefore, different agent-types and their corresponding subpopulations had to be implemented, in order to discriminate the different pathways possible for them to follow in the model.

The 724 234 patients that entered the waiting list in 2019 were considered to be part of a specific subpopulation of agents in the model, denominated *patients19* (which corresponding agent-type was *Patient19*). The remaining 244 501 patients, who were already on the waiting list at the start of the modelling year, were subdivided into three distinct groups, according to the stage of the waiting process they were in: *patients181* (agent-type *Patient181*) referred to the ones who, although already on the waiting list before 2019, had still not surpass 75% of their TMRG waiting and, therefore, were eligible to receive their first SV; *patients182* (agent-type *Patient182*) had waiting times between 75% and 100% of TMRG, meaning that they had already received their first SV and, in the developed model, were only entitled to receive a second SV (at 100% of TMRG); and, finally, *patients183* (agent-type *Patient183*) were the agents who, having surpassed 100% of TMRG waiting for surgery, were assumed to already having received the two SVs they have the right to be offered and, therefore, would not receive any transfer proposal during the simulation.

Regarding the subdivision of the total number of patients already enrolled on the list at the beginning of 2019 in the multiple populations created – *patients181*, *patients182*, and *patients183* –, no country-wide information was available. Nevertheless, data from a Portuguese public hospital was used to define the distribution of waiting patients between these three groups. According to the waiting list for surgery from that same hospital, at the end of 2018, 57% of the enrolled patients were in the situation of the *patients181* population, 13% belonged to the *patients182* group and, lastly, 30% corresponded to *patients183*'s cases. Thus, this was the distribution used in the AnyLogic model.

It is important to note that this software does not support such big numbers of agents, which imposed the use of approximated values. In order to facilitate the interpretation of results, it was determined that each agent in the model represented 1 000 patients in real-life, which means that the distribution of agents between the four classes of patients considered in the model was the following: 724 agents of the *Patient19* type, 139 of *Patient181*, 32 of *Patient182*, and 73 of *Patient183*.

Agents of all categories were characterized by seven parameters (Sex, Age, Region, Priority, PBr, CtA, ItA) and four variables ("NbSVs", "DidAcceptSV", "DidRefuseSV", "DidNotAnswerSV"). The distribution of the four demographic parameters was defined according to the description of the surgery waiting list population reported in a SIGIC's study [19]. For each patient, "NbSVs" denotes the number of SVs emitted, "DidAcceptSV" signals if an SV has been accepted or not, "DidRefuseSV" shows the number of SVs' refusals, and "DidNotAnswerSV" the number of SVs that did not obtain any answer by the patients they were issued to. PBr, CtA, and ItA represent each patient's behaviour constructs of interest – perceived barriers, cues to action and intention to accept. PBr and CtA values were set according to their respective indicators' scores observed in the HBM questionnaire sample. The mathematical expression defining ItA from PBr and CtA was implemented according to the results of the SEM's global model, after path analysis.

Although more than one agent-type were considered in order to represent patients in different circumstances of their waiting journey, the sequence of states and transitions such agents experienced was always based on adaptations of the same overall representation of the waiting list's universe – Figure 14 shows the statechart scheme which illustrates the functioning of the general model and, more specifically, the one followed by patients from the *patients19* subpopulation.

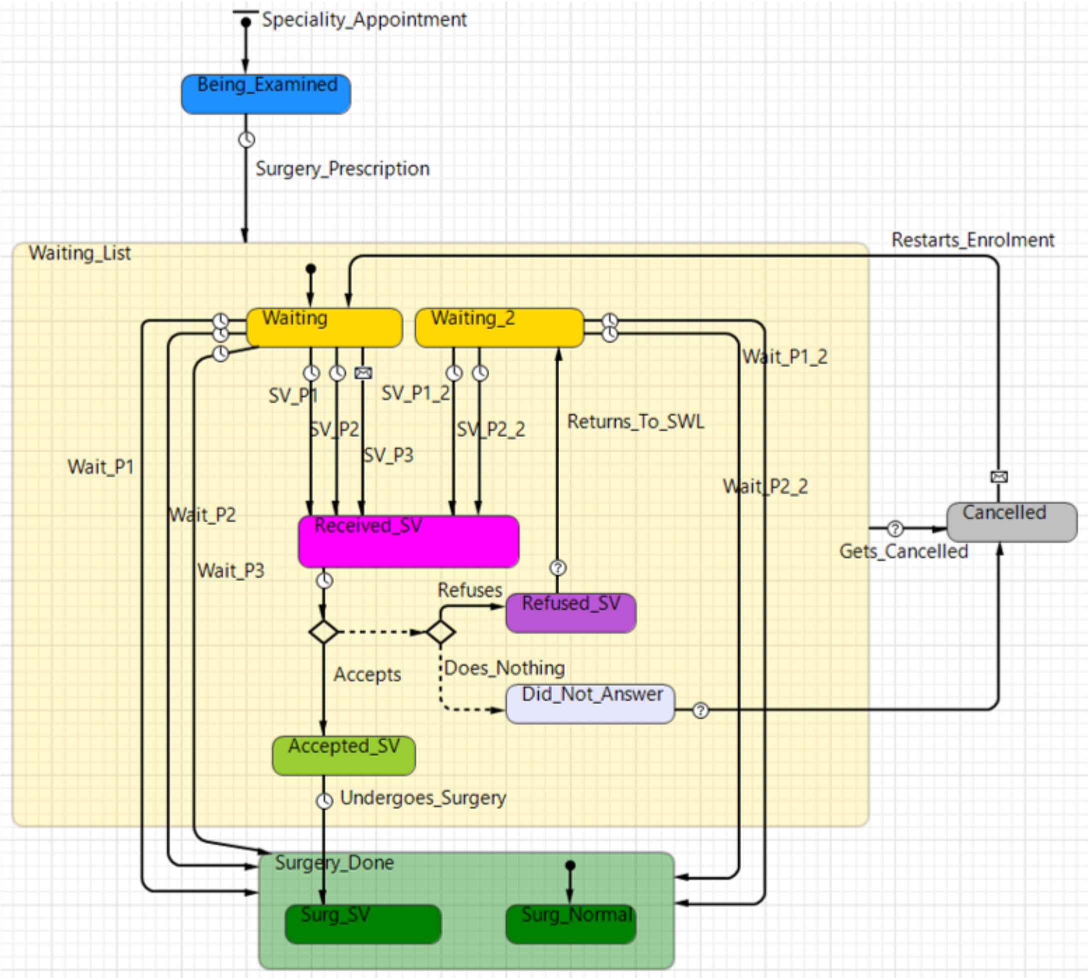

Figure 13: Statechart developed in AnyLogic to define agents' states and transitions in the simulation, as a representation of patients' possible pathways on the Portuguese surgery waiting list.

Agents enter the model environment at the "Speciality_Appointment" starting point, which symbolizes the first speciality medical appointment they go to. At that stage, begins the proposal phase of each patient's episode of care (explained in Chapter 2) – here represented by the "Being_Examined" state (in blue) –, which may encompass many medical encounters, and during which patients' care plan is designed. Considering this context, and for a matter of simplicity, there were only considered cases needing surgery, which means that all agents in the "Being_Examined" state will eventually be prescribed surgery and enrolled on the waiting list – represented by the "Surgery_Prescription" transition, leading to the "Waiting_List" superstate (in light yellow). This transition was implemented as a countdown based on a uniform distribution from zero to 365 days, so that all agents of the type *Patient19* entered the waiting list during the modelled year of 2019.

Once enrolled for surgery, patient agents enter the "Waiting_List" state, which includes multiple substates: "Waiting", "Waiting_2", "Received_SV", "Accepted_SV", "Refused_SV", and "Did_Not_Answer". In the beginning of the waiting process, the agent is automatically directed to the "Waiting" state, from where it can either proceed to “Received_SV” (in pink), in case that patient’s surgery does not get scheduled before the deadlines defined by law causing an SV to be emitted, or to “Surgery_Done” (in green), if surgery is performed within the expected time period.

What determines if a certain patient is going to directly undergo surgery or, instead, to receive an SV is a database table of true waiting times registered at a Portuguese public hospital for that specific patient’s priority level, which is defined in both transitions “Wait_P1” and “Wait_P2” (for priority levels one and two, respectively). On the other hand, transitions “SV_P1” and “SV_P2”, from “Waiting” to “Received_SV”, have a fixed countdown of 135 and 30 days, respectively and according to the information displayed in Table 1 regarding the timestamps for the emission of SVs and TNs. To each patient in “Waiting” is assigned a waiting time until scheduling extracted from the hospital’s database tables and identified by “Wait_P1” or “Wait_P2”. If that time is shorter than “SV_P1” or “SV_P2”, depending on the case, the patient is sent to the “Surgery_Done” state indicating that surgery was scheduled and, consequently, directly performed (more specifically, the patient goes to the “Surg_Normal” substate, which differentiates the interventions performed without the emission of SVs).

In the cases where the surgery waiting time (“Wait_P1” or “Wait_P2”) is longer than the time for emission of a transfer document (“SV_P1” or “SV_P2”), patients move to the “Received_SV” state, where their “NbSVs” count variable, which keeps track of the number of SVs emitted to a certain patient, is increased in one unit. There, agents’ destiny is dependent on their personal decision regarding the SV they were offered – they can either accept it and move to the “Accepted_SV” state, decline it and move to “Refused_SV”, or not answer, in which case they will be sent to the “Did_Not_Answer” state. The conditions on which patients proceed to one of these states are defined by the transitions “Accepts”, “Declines”, and “Does_Nothing”. As the subject of interest in this simulation study, “Accepts” was not fixed to a given percentage of acceptance, but rather implemented as a condition dependent on each agent’s ItA score. ItA above a specific threshold denoted that patients’ intention to accept an SV was high enough to originate that same behaviour, while ItA values lower than that limit would result in SVs’ refusals or in the absence of response. To define that cut-off value, the survey’s sample was used: participants’ ItA estimates were computed and the top 18.8% of those scores was considered to result in effective acceptance (in order to correspond to the reality of 2019). That revealed that values of intention to accept higher or equal to -4.39 would result in actual SVs’ acceptance, and this was the cut-off value used to differentiate patients’ intention to behave from true acceptance behaviour in the model. This allowed the testing of scenarios where the behavioural tendencies of patients regarding SVs were different and the observation of their impact on the waiting list’s results, as it will be explained further ahead. The remaining SV-related transitions, “Refused_SV” and “Did_Not_Answer”, were modelled as fixed parameters and also corresponded to percentages reported by the Ministry of Health in 2019: from the non-accepted vouchers, 67.2% were due to justified refusals and the remaining caused by an absence of answer from the patient [18].

From “Did_Not_Answer”, agents are moved to the “Cancelled” state representing the cancellation of patients’ enrolment on the waiting list, as it is predicted in SIGIC’s guidelines. After that, patients can be re-enrolled on the list by their hospital, only after a revaluation of their situation. That scenario is represented by the “Restarts_Enrolment” transition, triggered by a specific message (“re-enrolment”), that leads once again to the “Waiting” state, since these patients lose their antiquity on the list and their waiting time is reset to zero.

When a patient accepts the SV, he/she advances to the “Surgery_Done” state (in green), more specifically to “Surg_SV”, which indicates surgery was performed within the ambit of the SVs’ transfer programme. Between both states, there is a transition, “Undergoes_Surgery”, symbolizing the waiting period between the decision to accept the transfer and the performance of the surgery. No information is available regarding this specific time period; therefore, this time was arbitrarily defined as 30 days.

Finally, patients have the option to decline their transfer. When patients answer negatively, they are returned to their home hospital’s waiting list (“Returns_To_SWL”), without losing their previous position. Thus, this time they are not sent to “Waiting”, but to “Waiting_2” (in yellow). From there, the destiny options are similar to the ones previously explained for the “Waiting” state: depending on the waiting time assigned to patients, they can either follow “Wait_P1_2” or “Wait_P2_2” to “Surgery_Done” and have their surgery performed, or, on the other hand, follow “SV_P1_2” or “SV_P2_2” to “Received_SV”, where they would receive their second SV. The main difference here is the time definitions of all those four transitions: “Wait_P1_2” and “Wait_P2_2” are, once again, waiting time distributions based on a database from a Portuguese hospital; “SV_P1_2” and “SV_P2_2” are set to 100% of TMRG, i.e., 180 and 60 days.

Guards implemented as Java code in ”Received_SV” guarantee that each patient gets a maximum of two SV offerings, as it is defined in SIGIC’s regulation. The exceptional scenario where a third SV is emitted as a consequence of patients’ request was not modelled for the sake of simplicity and due to the lack of statistics referring to that particular situation.

At any moment patients’ enrolment on the waiting list can be cancelled (for a variety of reasons that include changes in the care plan or the death of the patient, for example). That corresponds to a transition from any part of the super-state “Waiting_List” to “Cancelled”, called “Gets_Cancelled”, defined as the percentage of enrolment cancellations experienced in 2019, 13.7% [18]. Once a patient’s enrolment is cancelled, he/she can be re-enrolled if the hospital decides so, as already explained, following the “Restarts_Enrolment” transition.

Patients with priority level three were only considered in the *patients19* subpopulation, due to their reduced TMRG. These patients only receive an SV in case they request so and, for that reason, the “SV_P3” transition was defined in the model to be triggered only upon a specific message (in this case, “request”). Despite this situation being considered in the model, it ended up not being actually applied in the simulation, since the sending process of such message was not implemented because of the lack of information regarding the amount of level three priority patients who ask for transfer.

As stated above, the four agent subpopulations were created due to patients’ pathways differences on the waiting list, according to the stage they were in at the beginning of the modelling year. Population *patients19* followed the statechart of Figure 14, since these patients started 2019 not yet enrolled on the list, while the remaining patient groups underwent slightly adapted versions of that sequence of states: *patients181* (statechart represented in Figure 21, Appendix B) start the simulation in the “Waiting” state, as they still have the opportunity to receive two SVs; *patients182* (statechart in Figure 22, Appendix B) begin at “Waiting_2”, since they already surpassed the timestamp for the first SV; and, lastly, *patients183* (statechart in Figure 23, Appendix B) also start the simulation process at “Waiting_2”, with the difference that both transitions “SV_P1_1” and “SV_P1_2” were removed from their statechart as they are not eligible to receive any more SVs.

Controls were implemented to vary behaviour determinants' estimates. PBr's control button enabled its values to fluctuate from 0% to 100%, and CtA's from 100% to 200%. In an initial run, they were both set to 100%, in order to achieve the baseline acceptance rate registered in 2019, 18.8% – scenario A. The outcomes of the model – number of SVs emitted, accepted, refused, and not answered, number of surgeries performed, both in total and through the emission of SVs, number of waiting list's enrolment cancellations, and number of patients enrolled on the waiting list – were used to understand if the model did satisfactorily reproduce the reality of that year, according to the Ministry of Health's reports. After that, six other simulation runs, where behavioural components PBr and CtA were modified (PBr decreased and CtA increased, so that acceptance would also grow), were executed, in order to evaluate its effects on the waiting list's outcomes: scenario B (PBr set to 75%, CtA set to 100%), scenario C (PBr set to 50%, CtA set to 100%), scenario D (PBr set to 100%, CtA set to 125%), scenario E (PBr set to 100%, CtA set to 150%), scenario F (PBr set to 75%, CtA set to 125%), and scenario G (PBr set to 50%, CtA set to 150%). Those results were then interpreted, to identify which behavioural interventions could have a bigger impact on patients' acceptance and on the Portuguese waiting list' universe.

4.4 Conclusions

To sum up, this behavioural research encompassed two main stages: modelling and simulation. Modelling started with the adaptation of HBM's constructs to the context of study – SVs' acceptance – and the establishment of seven research hypotheses regarding the associations depicted in the model. A measurement instrument, consisting on a questionnaire, was also developed, where each HBM construct was operationalized as a set of multiple survey's items. Answers to the questionnaire were collected and analysed through EFA, in order to understand if the answers' factor structure was similar to the HBM-inspired groups of items. Afterwards, CFA was used to confirm EFA's results. Cross-sampling was used in order to perform both factor analyses in independent subsamples. ItA was then added to the model, during path analysis, so that causal weights of each construct could be computed. That revealed which factors influenced behaviour, and to which extent.

The causal weights extracted from path analysis were added to individual agents' attributes in a simulation model, in order to infer behavioural consequences from intention to behave. Those agents represented patients on the Portuguese waiting list, enrolled for surgery. The model was developed based on the waiting process reported in SIGIC's guidelines and on waiting times from a Portuguese public hospital. Agents had their own personal characteristics, including HBM determinants of behaviour, and made their own decisions regarding SVs. Seven scenarios with differing behavioural trends were tested and their effects on the waiting list's outcomes compared.

All the obtained results will be presented and dissected in the following chapter.

Chapter 5

Results and Discussion

This chapter presents all the important results obtained. Subchapter 5.1 reffers to the characterization of the sample who answered the questionnaire; Subchapter 5.2 shows both EFA and SEM's results regarding the survey's responses, and discusses the validity of the proposed conceptual model along with its measurement instrument; finally, Subchapter 5.3 explores the simulation outcomes.

5.1 Characteristics of the sample

The developed questionnaire obtained a total of 177 responses. From those, only 170 were considered in the final sample – answers from residents in Azores and Madeira were excluded from the study. Table 5 shows the demographic characteristics of the considered sample, both in total terms and in relation to their intention to accept an SV. In addition, the table shows the demographic description of the sample used in a previous SIGIC's telephonic enquiry, which claimed to well-represent the Portuguese surgery waiting list's population. Therefore, it was possible to compare both samples and assess the representativeness of the one used in the present project.

Table 5: Sociodemographic characterization (age, gender, ARS, educational level) of both SIGIC's reference study's sample [19] and this dissertation's survey's sample – in total terms, as well as in relation to positive or negative intention to accept the SV. P-values of Fisher's independence test are also displayed.

Demographic variables		SIGIC's Telephone Enquiry		Present Study						
		Frequency (total=570)	Percentage (%)	Total Sample		"Intending to Accept"		"Not Intending to Accept"		Fisher's Test P-Value
				Frequency (total=170)	Percentage (%)	Frequency (total=157)	Percentage (%)	Frequency (total=13)	Percentage (%)	
Age (years)	0-20	45	7.9	3	1.8	3	1.9	0	0.0	0.306
	21-40	67	11.8	79	46.5	74	47.1	5	38.5	
	41-60	199	34.9	59	34.7	55	35.0	4	30.8	
	above 60	259	45.4	29	17.1	25	15.9	4	30.8	
Gender	Female	338	59.0	125	73.5	117	74.5	8	61.5	0.333
	Male	232	41.0	45	26.5	40	25.5	5	38.5	
ARS	North	157	27.5	5	2.9	5	3.2	0	0.0	1.000
	Centre	186	32.6	15	8.8	14	8.9	1	7.7	
	LTV	167	29.3	145	85.3	133	84.7	12	92.3	
	Alentejo	32	5.6	3	1.8	3	1.9	0	0.0.	
	Algarve	28	4.9	2	1.2	2	1.3	0	0.0	
Educational level	None	-	-	0	0.0	0	0.0	0	0.0	0.277
	Primary school	-	-	19	11.2	16	10.2	3	23.1	
	Middle school	-	-	7	4.1	6	3.8	1	7.7	
	High school	-	-	44	25.9	42	26.8	2	15.4	
	University	-	-	100	58.8	93	59.2	7	53.8	

According to Table 5, it is possible to verify that the sample considered in this study comprised an overrepresentation of young adults (21-40 years old) – 46.5% versus 11.8% obtained in the SIGIC enquiry – and an underrepresentation of the older age group (above 60) – only 17.1% versus 45.4%. Such results reflect the easiness to reach surveys' responses through online platforms, where young people are much more active than the elderly. As previously stated, door-to-door distribution of the questionnaire was used as a method to neutralize this tendency; however, it did not reach a sufficient countereffect, since it was quite difficult to get a high enough number of responses. The proportion of participants belonging to the 0-20 age group was slightly lower in the HBM questionnaire (1.8%), when compared with the reference study (7.9%). On the other hand, the representation of patients with ages between 41 and 60 years old was quite similar in the two enquiries, being close to reach 35%.

In both studies, there was a higher percentage of female over male participants (although that difference was more significant in the case of this dissertation). Moreover, it was difficult to engage with people from all geographical parts of the country, leading to a dominant representation of the LTV region (85.3%). The percentages of patients from Alentejo and Algarve were similarly low. However, both ARSs of North and Centre ended up being underrepresented in the present research.

Respondents' educational level was also assessed, since it was thought it could have some kind of influence in their decision regarding SVs. Almost 85% of the sample had, at least, completed high school. No comparisons could be made regarding educational information, as it was not assessed in SIGIC's reference study. It is also worth mentioning that this report made by the Ministry of Health dates from 2007, which means the corresponding characterization of the waiting list that here was used as a point of reference might as well be a little outdated.

In order to evaluate if any of these demographic variables played an important role on patients' ultimate decision to accept or not an SV, Fisher's test of independence was performed, in R. Results are also displayed in Table 5 for each sociodemographic characteristic queried. Every test reached a p-value higher than the predetermined alpha significance level (0.05), meaning that the null hypothesis (that the variables are independent) should not be rejected. Therefore, age, gender, ARS, and educational level were all considered independent from one's intention to accept an SV.

Besides demographic questions, the first part of the survey also included a ten-point item where participants had to rate their general health status (Table 6), as well as two questions related to patients' previous history on the surgery waiting list (Table 7). Regarding respondents' health self-evaluation, the Mann-Whitney U test produced a p-value of 0.544, indicating that, once again, answers to this item were independent from intention to accept.

About previous experience with SVs and the waiting list, more than 75% of the enquired had never been enrolled on the latter. From the 41 people who had been enrolled on the list for surgery, 23 never received an SV. From the 18 who received one, ten accepted it and used it, while eight did not use it. All ten participants who reported having used an SV in the past showed being willing to repeat the experience. To understand if these factors (having been previously enrolled on the waiting list, having previously received an SV, and having already used an SV) were independent from intention to accept an SV in the future, Fisher's test was applied. Previous enrolments on the waiting list proved to be

Table 6: Distribution of survey's answers regarding respondents' health status self-evaluation – in total terms, as well as in relation to positive or negative intention to accept the SV. The p-value of the Mann-Whitney U test is also displayed.

"In general, how do you evaluate your health status?"	Total Sample		"Intending to Accept"		"Not Intending to Accept"		Mann-Whitney U Test P-Value
	Frequency (total=170)	Percentage (%)	Frequency (total=157)	Percentage (%)	Frequency (total=13)	Percentage (%)	
1 (very poor)	0	0.0	0	0.0	0	0.0	0.544
2	1	0.6	1	0.6	0	0.0	
3	1	0.6	1	0.6	0	0.0	
4	5	2.9	5	3.2	0	0.0	
5	13	7.6	10	6.4	3	23.1	
6	13	7.6	13	8.3	0	0.0	
7	37	21.8	34	21.7	3	23.1	
8	46	27.1	42	26.8	4	30.8	
9	41	24.1	39	24.8	2	15.4	
10 (excellent)	13	7.6	12	7.6	1	7.7	

Table 7: Distribution of survey's answers regarding respondents' history on the surgery waiting list – in total terms, as well as in relation to positive or negative intention to accept the SV. The p-values of Fisher's independence tests are also displayed.

Surgery waiting list-related questions		Total Sample		"Intending to Accept"		"Not Intending to Accept"		Fisher's Test P-Value	
		Frequency (total=170)	Percentage (%)	Frequency (total=157)	Percentage (%)	Frequency (total=13)	Percentage (%)		
"Have you ever been enrolled on the surgery waiting list before?"	Yes	41	24.1	36	22.9	5	38.5	0.308	
	No	129	75.9	121	77.1	8	61.5		
"If yes, have you ever received an SV? Did you use it?"	I received it and used it	10	5.9	10	6.4	0	0.0	0.021	0.203
	I received it, but did not use it	8	4.7	5	3.2	3	23.1		
	I did not receive it	23	13.5	21	13.4	2	15.4		

independent from intention to accept a future SV (with a p-value of 0.308). The same happened when patients were divided into two different groups depending if they had received an SV somewhere in the past or not (with a p-value of 0.203). On the contrary, the fact that one had previously accepted or not an SV showed influence on the last question of the survey, presenting a significant p-value of 0.021, which suggested dependence between past and future decisions regarding SV's acceptance. This shows that, in spite of the low acceptance rates obtained among patients, the ones that took advantage of that transfer opportunity were not disappointed by the SVs system's functioning.

5.2 Validation of the SVs' acceptance model

This section includes the analysis of the survey items' factor structure, using EFA (Subchapter 5.2.1), the measurement model obtained during CFA (Subchapter 5.2.2), and SEM's global model (Subchapter 5.2.3). In the end, it presents a final evaluation of the research hypotheses previously established.

5.2.1 Factor structure (EFA)

Data from one of the two questionnaire's subsamples, comprising a total of 87 responses, was imported to R, so that EFA could be performed. No missing data or outliers were detected in the dataset. Table 8 shows some descriptive statistics regarding all 26 variables, from Q1 to Q26. For example, it is possible to observe that items' mean values varied from 1.89 to 4.60. Skewness was always under $|2.0|$ with the exception of item Q2, where it reached -2.15, while kurtosis was inferior to $|2.0|$ in 23 out of the 26 variables and ranged from $|2.0|$ to $|6.0|$ in the remaining four.

Table 8: Statistics obtained for questionnaire variables Q1 to Q26: mean value, standard deviation, median, skew, and kurtosis.

Variable	Mean	Standard Deviation	Median	Skew	Kurtosis
Q1	4.43	0.86	5	-1.68	2.79
Q2	4.56	0.74	5	-2.15	5.80
Q3	4.47	0.79	5	-1.72	3.51
Q4	4.39	0.70	5	-0.90	0.24
Q5	4.21	0.82	4	-0.64	-0.58
Q6	4.60	0.56	5	-0.97	-0.10
Q7	4.22	0.77	4	-0.69	-0.11
Q8	4.25	0.77	4	-0.76	0.02
Q9	4.29	0.82	4	-0.94	0.14
Q10	2.97	1.37	3	0.03	-1.25
Q11	2.93	1.29	3	0.00	-1.14
Q12	3.18	1.26	3	-0.07	-1.15
Q13	3.15	1.36	3	-0.13	-1.22
Q14	2.21	1.17	2	0.84	-0.19
Q15	1.89	1.10	2	1.30	0.99
Q16	3.57	1.28	4	-0.75	-0.52
Q17	3.15	1.23	3	-0.21	-0.85
Q18	4.08	1.03	4	-1.18	1.02
Q19	3.82	1.04	4	-0.68	-0.25
Q20	4.45	0.87	5	-1.92	4.04
Q21	4.05	0.91	4	-0.54	-0.73
Q22	4.15	1.03	4	-1.25	0.89
Q23	4.26	0.74	4	-0.45	-1.08
Q24	3.61	1.03	4	-0.32	-0.57
Q25	4.18	0.67	4	-0.23	-0.86
Q26	4.06	0.91	4	-0.67	-0.42

Although such values of skewness and kurtosis do not necessarily imply a significant absence of normality in variables' distributions, the ordinal nature of the data here considered suggests that scenario should be carefully taken into account, since it may influence the type of correlation coefficients that should be used to conduct this analysis. Thus, Mardia's multivariate normality test was performed and statistically significant skewness and kurtosis values ($p<0.05$) were obtained. That implied the rejection of the null hypothesis, stating that variables followed a multivariate normal distribution, and led to the conclusion that computing polychoric correlation coefficients (instead of Pearson coefficients, for example) was truly the most suitable approach to follow. The plot of the resultant polychoric correlation coefficients among variables Q1 to Q26 is shown below, in Appendix C Figure 24.

To assess the factorability of the correlation matrix from Figure 12, both the KMO and Bartlett's tests were performed. The obtained model sampling adequacy through the KMO test was 0.68, higher than the established minimum of 0.50, and Bartlett's test of sphericity reached high statistical significance ($p<0.05$), indicating that the correlation matrix was not random. Therefore, it was considered adequate to carry on with factor analysis.

Common factor analysis was performed in order to explore the latent structure underlying HBM-related items' answers, with principal axis factoring being the chosen mathematical method for factor extraction. Parallel analysis was the main criterion used to determine the suitable number of factors to be considered in further analysis. While the theory where the survey was based on suggested seven underlying factors (corresponding to the HBM latent variables: perceived susceptibility, perceived severity, perceived benefits, perceived barriers, cues to action, self-efficacy, and health motivation), parallel analysis indicated the presence of only five factors. A scree plot examination corroborated the five-factor answer, as it is shown in Figure 14, where the fifth factor was also an inflexion point, and the five initial factors were all above their corresponding randomly generated eigenvalues (represented by the red dotted line), as well as above the “eigenvalue one” black line). Since multiple criteria seemed to agree that five was the suitable number of factors to retain, there was no need to consider multiple hypotheses, and the analysis continued with the established five constructs (once again using principal axis factoring, and *oblimin* as the oblique rotation method).

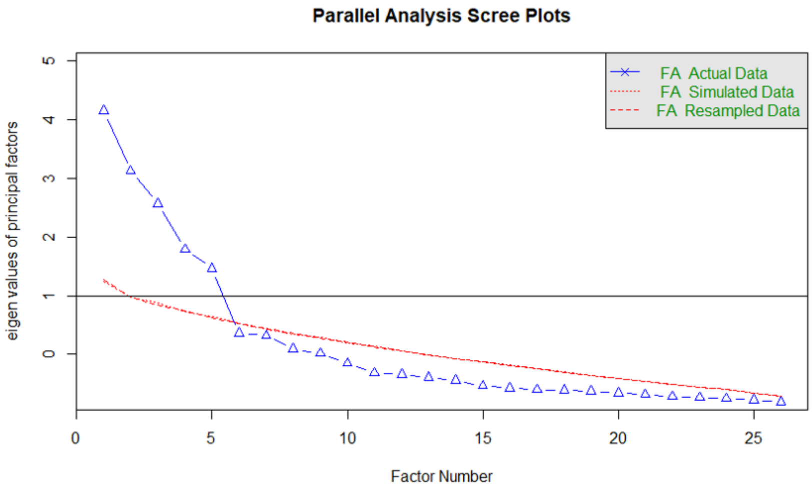

Figure 14: Parallel analysis scree plots, representing factors’ eigenvalues as a function of the number of factors (ordered in descending order of their corresponding eigenvalues). The blue line (also marked with triangles), ”FA Actual Data”, represents the factor analysis of the sample data; the red dotted line, ”FA Simulated Data”, signifies a randomly obtained factor-eigenvalue distribution; the red dashed line, ”FA Resampled Data”, represents a resampling of the original data. The number of factors in blue (extracted from the questionnaire data) located above the intersection of the curves can be used as a clue to the ideal number of factors to be considered in further analysis. Kaiser’s criterion can also be used – here represented by the black line, signalling an eigenvalue of one.

Rotated loading factors obtained for all 26 measured variables are displayed in Table 9, where significant values (above 0.5) were highlighted in bold. No noteworthy cross-loadings were identified. However, items Q21 and Q22 did not show acceptable loadings in any of the considered factors. These questions, which theoretically seemed to well-represent self-efficacy, empirically demonstrated to not operationalize any meaningful construct. That could be explained by the fact that, ideally, a latent variable should be represented by, at least, three observable indicators. However, in this case, these two specific items were thought to satisfactorily cover all aspects of their corresponding construct. Practical results indicate otherwise – by using only two measured variables, covariance was probably not significant enough to differentiate self-efficacy as a considerably important dimension of the model. Furthermore, both indicators were also candidates to exclusion when communalities were assessed, with both values being lower than 0.40.

Even when mathematical criteria point to the elimination of certain survey’s items, it is possible to justify their inclusion in a study if theoretical arguments support that decision. However, in this case-

Table 9: Observed variables (Q1 to Q26), their corresponding factor loading in all five considered factors (PA1 to PA5), as well as their communality and uniqueness values. Factor loading coefficients higher than the 0.50 cut-off are written in bold.

Variable	Factor Loading Coefficients (λ_{ij})					Communality (h^2)	Uniqueness (u^2)
	PA1	PA2	PA3	PA4	PA5		
Q1	0.04	**0.56**	-0.13	-0.07	0.01	0.33	0.67
Q2	-0.04	**0.71**	0.01	-0.03	-0.09	0.49	0.51
Q3	0.01	**0.89**	-0.01	-0.09	-0.08	0.76	0.24
Q4	0.01	**0.78**	0.05	0.04	0.08	0.64	0.36
Q5	-0.04	**0.65**	-0.01	0.26	0.13	0.60	0.40
Q6	-0.08	**0.63**	0.08	0.05	0.20	0.52	0.48
Q7	0.03	0.02	-0.07	-0.11	**0.85**	0.76	0.24
Q8	0.09	-0.01	-0.02	0.04	**0.93**	0.86	0.14
Q9	-0.05	0.05	-0.01	-0.03	**0.87**	0.78	0.22
Q10	**0.74**	-0.15	0.08	-0.05	-0.05	0.62	0.38
Q11	**0.79**	-0.13	0.02	-0.02	-0.06	0.67	0.33
Q12	**0.88**	0.04	-0.03	0.01	0.05	0.77	0.23
Q13	**0.91**	0.02	-0.01	0.05	0.03	0.82	0.18
Q14	**0.57**	0.15	0.03	0.08	0.18	0.37	0.63
Q15	**0.60**	0.12	0.09	0.08	0.07	0.39	0.61
Q16	0.13	0.10	**0.85**	-0.09	-0.14	0.83	0.17
Q17	0.07	0.11	**0.77**	0.00	-0.05	0.64	0.36
Q18	-0.04	-0.15	**0.82**	0.03	0.03	0.69	0.31
Q19	-0.12	0.00	**0.77**	0.06	0.02	0.59	0.41
Q20	-0.01	-0.08	**0.55**	0.02	0.21	0.31	0.69
Q21	-0.29	-0.05	0.14	0.44	0.19	0.34	0.66
Q22	-0.34	0.00	0.05	0.41	0.28	0.40	0.60
Q23	0.02	0.07	-0.02	**0.89**	-0.03	0.81	0.19
Q24	0.03	0.06	-0.12	**0.65**	-0.27	0.46	0.54
Q25	0.07	-0.14	0.08	**0.70**	-0.04	0.50	0.50
Q26	0.05	-0.03	0.00	**0.72**	0.04	0.53	0.47

study, self-efficacy seems to not crucially influence SVs' acceptance even in theoretical terms, since the health-related behaviour here considered concerns a personal medical decision, that does not require any special kind of skill or capacity (either psychological, physical, or technical) to be executed, which may otherwise happen in different situations. Examples of other contexts where self-efficacy was significantly important are Taiwan's personal health book [12], where one's familiarity with information technologies could influence the use of such tool, and smoking cessation studies [54], where smokers' perceptions regarding their own capacity to quit that habit also played an important role in eventual behavioural changes. Moreover, self-efficacy is not always present in HBM studies, since it was a later add-on to the model (because of its value in situations like the ones just mentioned) that, however, does not necessarily fit in every context. In conclusion, there seemed to be no reason to maintain both items Q21 and Q22 in this study and, therefore, they were excluded from the analysis. The removal of the self-efficacy dimension from the model also meant that research hypothesis H7 ("Self-efficacy is positively associated with intention to accept an SV") was not supported by the model.

Each one of the other variables presented high loading values on one and only one factor and were, therefore, kept in the analysis. New factor loading estimations for the remaining 24 variables were obtained after the exclusion of items Q21 and Q22 from the dataset. No significant differences were observed for the lasting variables, whose new loadings are shown in Appendix C, Table 17.

As it is observable, factor PA1 is saliently loaded by items Q10, Q11, Q12, Q13, Q14, and Q15, which were intended to correspond to the perceived barriers HBM construct. Thus, this factor was considered to represent such latent variable. Items Q1, Q2, and Q3 were created to indirectly measure perceived severity, while Q4, Q5, and Q6 were initially envisioned to represent perceived susceptibility. However, during EFA, all six items were collapsed into a same factor, PA2. This union is supported by theory, since the HBM model considers that perceived susceptibility and perceived severity together form a super-construct called perceived threat, as it is represented in Figure 5. Variables Q16 to Q21 were, as expected, grouped into a same factor, PA3, which was then named cues to action, after the latent variable it stands for. Questions Q23 to Q26 corresponded to PA4 and concerned health motivation. Finally, factor PA5, formed by Q7, Q8, and Q9, represented perceived benefits of accepting an SV. Table 10 summarizes the correspondence between these factors, their questionnaire items, and the HBM latent variables they stand for, as well as each factor's proportion variance. Factor PA1 is the one responsible for a higher total variance of the collected survey's answers, corresponding to 15%. PA2 explains 14% of the total variance, PA3 12%, PA5 11%, and lastly PA4 with 10%..

Table 10: Factors' corresponding HBM constructs, the amount of answers' variance they are responsible for, and, finally, their respective Chronbach's alpha values.

Factor	HBM Construct	Variance	Cronbach's α
PA1	Perceived Barriers (PBarr)	0.15	0.89
PA2	Perceived Threat (PThreat)	0.14	0.86
PA3	Cues to Action (CtA)	0.12	0.86
PA4	Health Motivation (HM)	0.10	0.82
PA5	Perceived Benefits (PBen)	0.11	0.93

Regarding correlations among factors, the majority was inferior to $|0.10|$. The most significant correlations happened between PA2 and PA5 (0.21), PA3 and PA5 (-0.18), and PA1 and PA3 (0.16). All factors obtained high Cronbach's alpha values, which are also specified in Table 10. Namely, PA5 reached 0.93, PA2 and PA3 both 0.86, and PA4 0.82 – all very good reliability coefficients that point to good internal consistency among the items that composed each factor.

To wrap up the exploratory part of the study, it is possible to conclude that EFA exposed an HBM-similar latent structure from the answers collected with the questionnaire developed within this dissertation. The above-presented results show that a five-factor solution is adequate to represent the SVs' HBM-based model, at least for the subsample being used. This solution had a very similar structure to the seven-construct conceptual model described in Subchapter 4.1, with the exception of self-efficacy's elimination, as already discussed and justified. Therefore, the measurement instrument proved to be adequate and valid to evaluate SVs' acceptance for the considered set of data. In order to infer about the robustness of this instrument and its corresponding model when a different dataset is considered, the following step was to perform CFA with the remaining subsample of survey's answers.

5.2.2 Measurement model (CFA)

The remaining half of the survey's sample (made out of 83 subjects) was imported to R and Mardia's multivariate normality test was performed. Statistically significant results for both skewness and kurtosis were obtained, corroborating the adequacy of using polychoric correlation coefficients to conduct CFA.

The measurement model, built in R, is displayed in Figure 15. The results confirm that the 24 considered indicators represent satisfactorily well the structure of five factors exposed during EFA. Nevertheless, some items showed standardized factor loadings below the cut-off of 0.50 and, for that reason, were eliminated at this stage of the study. That was the case of Q14 and Q15, which were poorly loaded in the perceived barriers factor. Perceived threat, perceived benefits, cues to action and health motivation did not suffer any modifications, since all indicators presented factor loadings higher than 0.50 concerning their specific constructs.

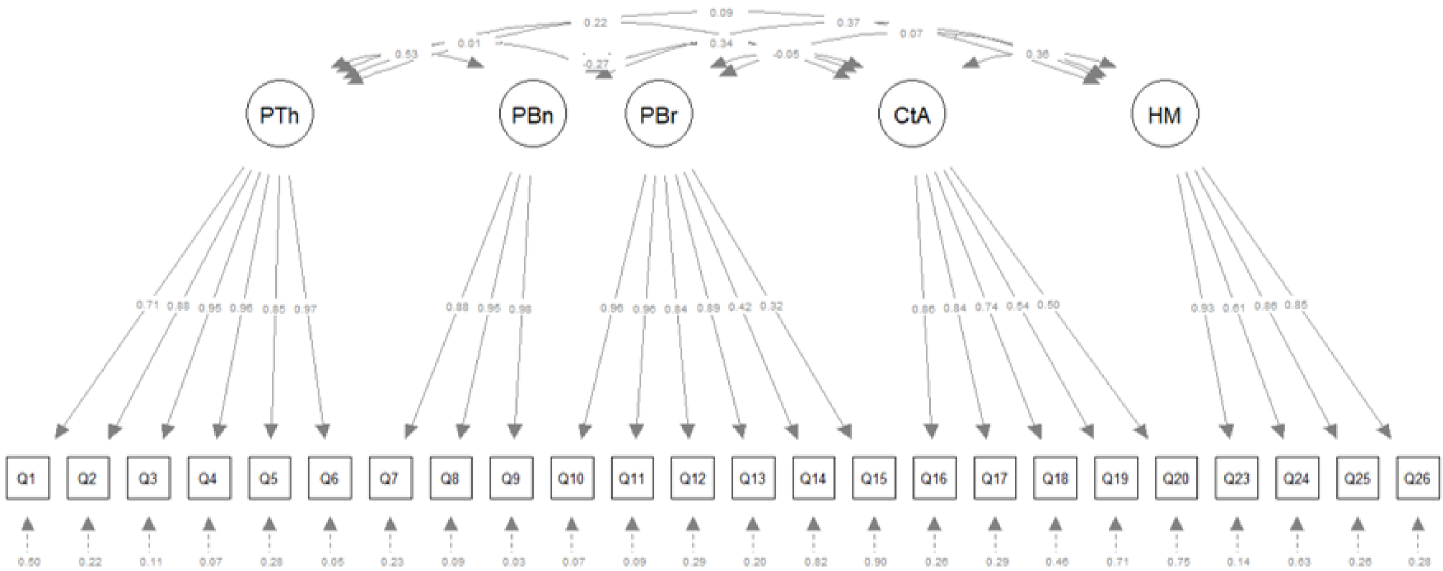

Figure 15: Measurement model obtained during CFA, considering 24 observed variables and five factors (PTh as perceived threat, PBn as perceived benefits, PBr as perceived barriers, CtA as cues to action, and HM as health motivation). Correlations among factors (ϕ_{jj}) are represented as double-ended arrows, while factor loadings (λ_{ij}) are directed arrows from each factor to their corresponding indicators. Indicators' error components (δ_i) are represented at the bottom, as dotted arrows pointing to each measured variable's square.

CFA was redone with the 22 lasting variables. Figure 16 presents the obtained results, where it is possible to observe indicators' factor loadings, correlations among factors, as well as the error components associated to each indicator's measurement (R^2 values can be obtained based on these error estimates, using equation (5)). All items showed standardized factor loadings higher or equal to the cut-off criterion, suggesting factor validity. Moreover, items showed satisfactory individual validity: all items exhibited λ_{ij}^2 values of at least 0.25, and the majority of them also presented R^2 estimates higher than 0.50, meaning that a major part of their variance was explained by their corresponding factor in the model. However, some exceptions were detected: indicators Q19, Q20, and Q24 obtained lower determination coefficients, of 0.293, 0.248, and 0.371, respectively. These items were, nonetheless, satisfactorily correlated to their corresponding factors – cues to action (Q19 and Q20) and health motivation (Q24) –, with loading coefficients equal to or slightly over the exclusion threshold of 0.50 (see

Figure 16). Since factor validity was not compromised, they were all kept in the analysis.

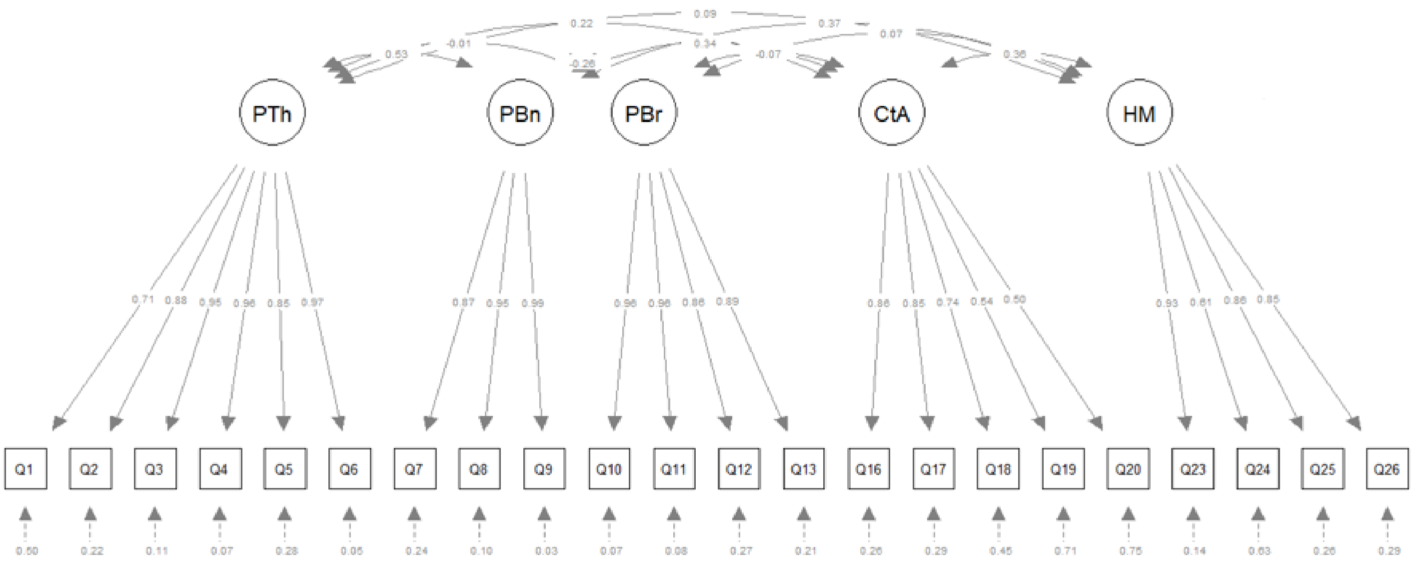

Figure 16: Measurement model obtained during CFA, considering 22 observed variables and five factors (PTh as perceived threat, PBn as perceived benefits, PBr as perceived barriers, CtA as cues to action, and HM as health motivation). Correlations among factors (ϕ_{jj}) are represented as double-ended arrows, while factor loadings (λ_{ij}) are directed arrows from each factor to their corresponding indicators. Indicators' error components (δ_i) are represented at the bottom, as dotted arrows pointing to each measured variable's square.

Correlations between factors were found to be higher than the ones previously detected during EFA. As expected from the previous analysis, perceived threat and benefits were the two constructs with higher correlation (0.53), followed by perceived benefits and health motivation (0.37), cues to action and health motivation (0.36), perceived benefits and cues to action (0.34), and perceived threats and cues to action (0.22). Perceived benefits and barriers showed a negative correlation (-0.26). The remaining values were all below $|0.09|$.

Regarding model's fit quality, extracted estimations of the evaluation indices mentioned in Subchapter 4.2.3 are displayed in Table 11. Overall, the model of Figure 16 proved to have a good fit to the studied dataset, which means it has a good capacity to reproduce the correlation structure observed in the measured sample [56].

Table 11: Estimated results for the goodness-of-fit tests applied to the measurement model (Figure 16) and their consequent qualitative evaluation.

Test	Estimate	Fit Evaluation
X^2/df	1.463	Good
CFI	0.994	Very good
TLI	0.993	Very good
RMSEA	0.075	Good

Factors' CR estimates were computed and are displayed in Table 12. As one may observe, they are all above 0.70, which indicates a good composite reliability of each construct.

Since factor validity seemed to be assured (by items' factor loading coefficients and internal validity), factors' AVE estimates were computed (Table 13), in order to evaluate other types of validity. All AVE

Table 12: Factors' composite reliability (CR) estimates.

Factor	CR
Perceived Threat	0.96
Perceived Benefits	0.96
Perceived Barriers	0.96
Cues to Action	0.83
Health Motivation	0.89

values were higher than 0.50, therefore suggesting the presence of convergent validity. In addition, for every factor, its AVE was considerably higher than the square values of its correlations with any other factor, r_{ij}^2, as it may also be observed in Table 13. To conclude, the developed model also presented discriminant validity.

Table 13: Comparison of each factor's AVE value (in bold) and its correlation square values with all other factors.

	PTh	PBn	PBr	CtA	HM
PTh	**0.73**				
PBn	0.29	**0.89**			
PBr	0.00	0.07	**0.67**		
CtA	0.05	0.11	0.01	**0.60**	
HM	0.01	0.14	0.00	0.13	**0.69**

To conclude, this CFA corroborated the structure unravelled by EFA of correlations between HBM's indicators and constructs, with the exception of two items which did not produce statistically significant correlations to remain in the model and were, therefore, removed. The model achieved a good adjustment to the data being studied and showed sound proofs of both reliability and validity – assuring that the final considered questionnaire items did indeed measure the HBM constructs they were related to in the measurement model, and did so in a consistent way.

5.2.3 Global model (Path Analysis)

The measurement model of Figure 16 was adapted to include the causal relationships among latent variables, which constitute the structural model. Therefore, the union of both models resulted in the construction of a SEM global model, which is shown in Figure 17. This model contains, besides the five constructs already explored during CFA, the behavioural dimension of the HBM model – intention to accept –, measured by item Q27.

The global model presented a good fit to the data, as it can be confirmed based on the quality adjustment tests' results, exhibited in Table 14. Thus, the measurement instrument developed in this dissertation, and operationalized through the model displayed in Figure 17, proved to be adequate to measure and predict SVs' acceptance behaviour.

Path analysis revealed that "Perceived barriers -> Intention to accept" was the most significant trajectory of the model, with a standardized weight of -0.63 (p=0.000). "Cues to action -> Intention to accept" came in second place, with a trajectorial weight of 0.19 (p=0.025). All the other three constructs

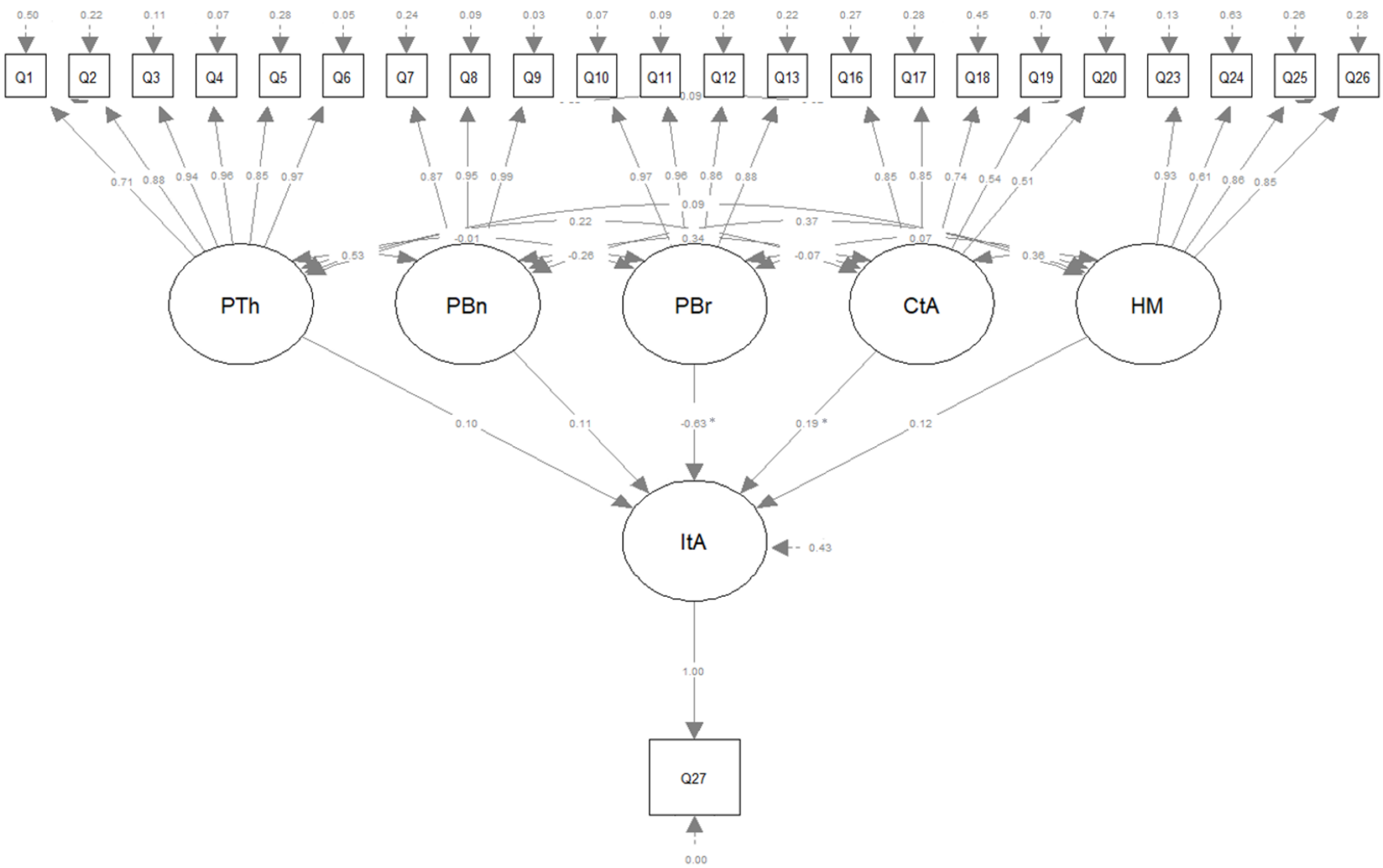

Figure 17: Global model obtained during path analysis, containing the five causal factors (PTh as perceived threat, PBn as perceived benefits, PBr as perceived barriers, CtA as cues to action, and HM as health motivation) as well as the behaviour latent variable (ItA as intention to accept). Correlations among factors (ϕ_{jj}) are represented as double-ended arrows, while factor loadings (λ_{ij}) are directed arrows from each factor to their corresponding indicators. Directed arrows from the five independent constructs to ItA are represented along with their respective path coefficients (γ_j), signalled with an * when statistical significance was achieved (p-values lower than 0.05). Indicators' error components (δ_i) are represented as dotted arrows pointing to each measured variable's square.

Table 14: Estimated results for the goodness-of-fit tests applied to the global model (Figure 17) and their consequent qualitative evaluation.

Test	Estimate	Fit Evaluation
X^2/df	1.368	Good
CFI	0.995	Very good
TLI	0.994	Very good
RMSEA	0.067	Good

(perceived threat, perceived benefits, and health motivation) did not obtain significant causal weights in relation to intention to accept: 0.10 (p=0.179), 0.10 (p=0.346), and 0.12 (p=0.126), respectively.

Therefore, regarding the six research hypotheses still under study, only H4 ("Perceived barriers are negatively associated to intention to accept") and H5 ("Cues to action are positively associated to intention to accept") could be accepted. Although perceived threat, perceived benefits, and health motivation had reached positive causal weights in relation to intention to accept, as hypothesised, such values did not show statistical significance and, because of that, H1, H2, H3, and H7 could not be accepted. Table

15 summarizes the final evaluation of all seven hypotheses.

Table 15: Research hypotheses (H1 to H7) under study and their final evaluation (accepted/not accepted).

Research Hypotheses	Evaluation
H1: Perceived susceptibility is positively associated to intention to accept an SV	Not accepted
H2: Perceived severity is positively associated to intention to accept an SV	Not accepted
H3: Perceived benefits are positively associated to intention to accept an SV	Not accepted
H4: Perceived barriers are negatively associated to intention to accept an SV	Accepted
H5: Cues to action are positively associated to intention to accept an SV	Accepted
H6: Self-efficacy is positively associated to intention to accept an SV	Not accepted
H7: Health motivation is positively associated to intention to accept an SV	Not accepted

The fact that perceived barriers significantly influenced patients' intention to accept an SV corroborated the conclusions obtained in a SIGIC's study [19], where some of those barriers were pointed by patients themselves as the main reasons for SVs refusals. Furthermore, incentives from people whose opinions patients value, such as family, friends, or their usual doctor, as well as a deterioration of their health condition while waiting for surgery – which together form cues to action – seemed to also have a somewhat important impact on their final decision to use an SV.

An analysis of the answers' distribution relative to the last questionnaire item (Figure 18), Q27, shows that the big majority of participants had a neutral (35%) or positive (57%) intention regarding SVs. Therefore, it is reasonable to infer that patients must be, in general and a priori, relatively prone to accept an SV. Nonetheless, it is known that, in practice, that is not translated in a satisfactory acceptance rate.

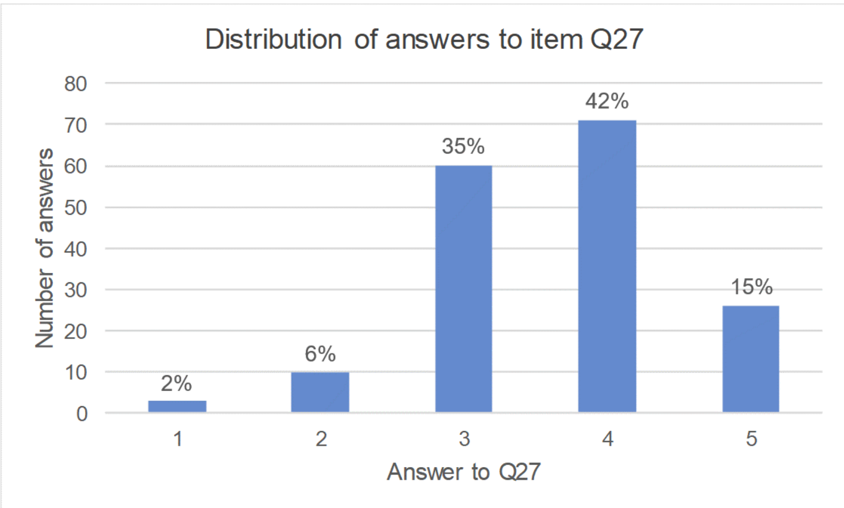

Figure 18: Distribution of the answers to the last item of the questionnaire, Q27, regarding patients' a priori willingness to accept an SV offer, considering a five-point Likert scale (as detailed in Table 4).

Thus, despite having a tendentially positive predisposition towards SVs, some aspects of the process lead the majority of patients to reject their transfer when presented with that possibility. In the model, the obstacles that emerge associated to the transfer process (measured by perceived barriers) proved to have an important negative effect on intention to behave. That and the fact that such barriers have been reported as the main reasons for SVs' rejection suggest that this HBM construct must be determinant to effective acceptance behaviour and the main cause of SVs' system's overall low performance. In addition, the SEM model showed that cues to action also partly explain patients' final decision, this time with a positive influence on intention to accept. Therefore, both constructs must be considered when developing policies aimed at increasing SVs' results, as it will be further explored next.

As explained in the methodology chapter, causal weights from SEM's global model for the statistically significant behavioural factors, perceived barriers and cues to action, were used in order to mathematically explain intention to accept, as follows:

$$ItA = -0.63 * PBr + 0.19 * CtA, \tag{11}$$

expression that was then inserted in the simulation model and that contributed to the results exposed in the following subchapter.

5.3 Simulation model

In order to understand if the results obtained in the simulation model regarding the surgery waiting list's evolution for scenarios B to G were reliable, it was mandatory to start with an evaluation of scenario A (where patients' behavioural determinants, PBr and CtA, were set to values congruent with the acceptance level registered in Portugal during 2019, 18.8%). It is, therefore, expected that the outcomes obtained for scenario A over the 365 days of the simulation running would be close to what occurred in reality. Table 16 establishes such comparison.

Table 16: Surgery waiting list's statistics in reality (2019) and in the baseline simulated scenario (scenario A). The relative errors concerning the real and simulated estimates are also represented.

	2019	Scenario A	Relative Error
Nb. of patients already in the SWL	244 501	244	0.2%
Nb. of patients entering the SWL	724 234	724	0.3%
Nb. of SVs emitted	249 962	276	10.4%
Nb. of SVs accepted	46 992	52	10.7%
Nb. of total surgeries performed	628 282	625	0.5%
Nb. of surgeries performed with SVs	-	45	-
Nb. of cancelled enrolments in the SWL	99 707	115	15.3%
Nb. of patients leaving the SWL	727 795	740	1.7%
Nb. of patients remaining in the SWL	242 949	227	6.6%

At the one-year mark of the simulation running, a total of 276 SVs had been emitted – which represent approximately 276 thousand SVs in real-life – and only 52 of those were accepted, corresponding to 18.8%. The value reported by the Ministry of Health is slightly lower, of 249 962 issued SVs, corresponding to an estimation error of about 10%. In the model, 625 (thousand) surgeries were performed, from which 45 were a result of the SVs' programme. The real total number of surgeries executed in 2019 was quite close, 628 282. Finally, summing up surgeries and cancellations unravels the number of patients who exited the waiting list – 740 (thousand) cases in the simulation versus 727 795 in reality, corresponding to an error close to 2%. At the end of the 365 days, 227 (thousand) were the patients still enrolled and that, for that reason, would transit to the 2020's waiting list. In real-life, that value was slightly higher, 242 949 (an error of 6.6%). Since the values obtained in the simulation seemed acceptably close to the ones observed in reality, the surgery waiting list's model here developed was considered adequate to study the effects of changes in patients' behaviour regarding a transfer offer.

Figure 19 shows the changes in SVs' general acceptance percentage among patients, detected for scenarios B (where PBr were decreased in 25%), C (where PBr dropped 50%), D (where CtA increased 25%), E (where CtA rose 50%), F (where scenarios B and D were combined) and G (a combination of C and E). As expected from the quantitative study's results, variations in perceived barriers (scenarios B and C) had a much stronger impact in intention to accept than modifications of the cues to action construct (D and E). In fact, while a 25% decrease in PBr caused acceptance to more than duplicate, an equivalent growing in CtA did not have a significant impact on the considered output. It is also important to mention the impressive value of 75.5% of acceptances obtained for a 50% reduction in PBr. A combined effect of both constructs, as it is possible to observe for scenarios F and G, resulted in an even higher impact. Nevertheless, PBr's much stronger contribution to changes in patients' decision-making is quite clear – thus, it should be the main target of health authorities when trying to combat SVs' declines.

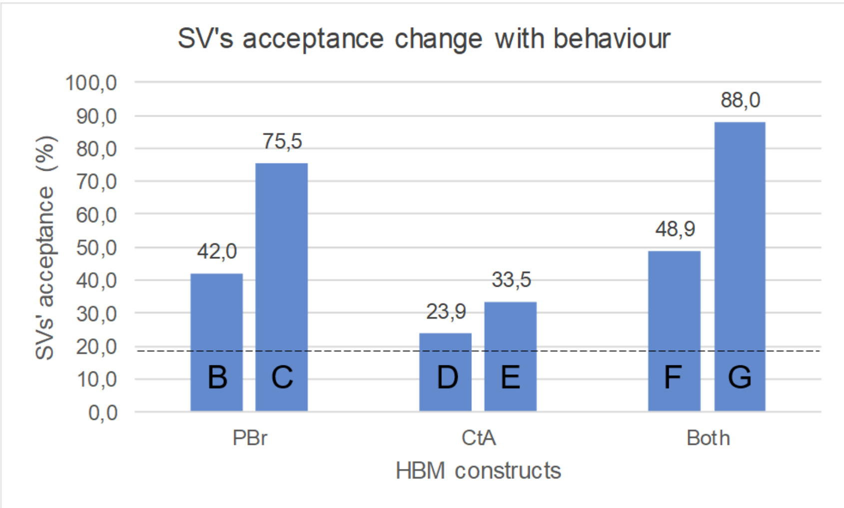

Figure 19: SVs' acceptance percentage for simulation scenarios B to G. The black dashed line marks the reference acceptance percentage, registered for scenario A.

To examine the effect of the acceptance levels depicted in Figure 19 on the national surgery delivery capacity, the graph of Figure 20 was built. With the exception of scenario D, that obtained results similar to scenario A, all the others presented a considerable increase in surgery delivery. In particular, scenarios E and F obtained similar results, with 687 and 690 surgeries performed – which would correspond to approximately 687 000 and 690 000 in reality. In scenario C, 720 surgical procedures were conducted (more 95 than in A). Scenario G, as expected, reached an even more impressive improvement, with 110 more surgeries than in the reference case. Once again, these values would have to be converted to thousands in order to foresee their possible impact in real-life.

Extending the scope of this impact analysis, Table 17 condenses the overall effects of the six behaviour settings on other metrics of the surgery waiting list, besides the absolute number of performed

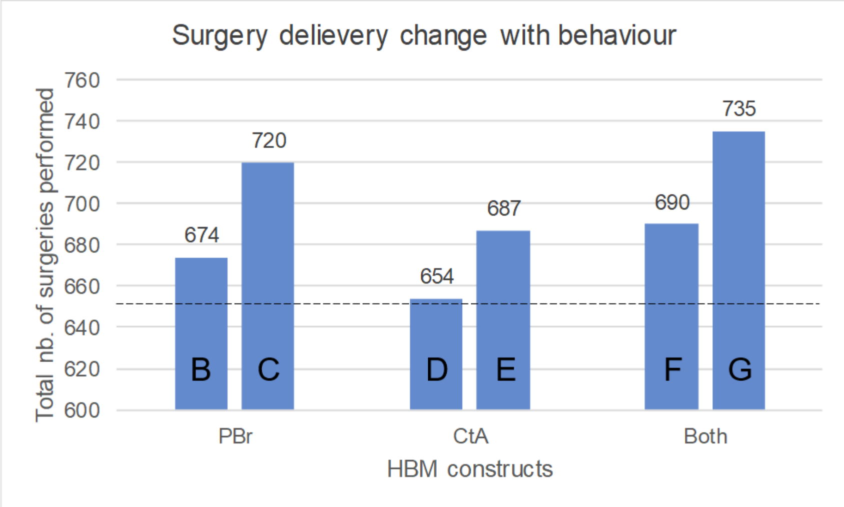

Figure 20: Total number of surgeries performed for simulation scenarios B to G. The black dashed line marks the reference number of surgeries, registered for scenario A.

surgeries already examined above. This figure shows a comparison between scenarios B to G, and the baseline scenario, A, the 2019-equivalent simulation, illustrating how those numbers would have been affected if behavioural transformations concerning patients' perceived barriers and cues to action estimates had occurred in that same year.

Table 17: Surgery waiting list's statistics for scenarios A to G.

	Scenario A	Scenario B	Scenario C	Scenario D	Scenario E	Scenario F	Scenario G
Nb. of patients already in the SWL	244	244	244	244	244	244	244
Nb. of patients entering the SWL	724	724	724	724	724	724	724
Nb. of SVs emitted	276	270	265	268	263	274	292
Nb. of SVs accepted	52	116	200	64	88	134	257
Nb. of total surgeries performed	625	674	720	654	687	690	735
Nb. of surgeries performed with SVs	45	104	181	60	80	118	223
Nb. of cancelled enrolments in the SWL	115	85	52	102	71	76	45
Nb. of patients leaving the SWL	740	759	772	756	758	766	780
Nb. of patients remaining in the SWL	227	208	195	211	209	201	187

For example, in scenario B, the model issued 270 vouchers – 6 less than in scenario A. From those, 116 (42.0%) received a positive response from the patients they were issued to, 64 more than in A. Moreover, a total of 674 surgeries were conducted, 104 of those through the emission of SVs. That represents an overall increase of 49 surgical procedures executed and a growth in SIGIC's impact of 59 more surgeries – 59 000 if converted to real-life impact. At the end of the year, 19 less patients were waiting for surgery, when compared to the final result of scenario A. Therefore, in a scenario where public health policymakers are able to increase patients' general acceptance of SVs to 42%, it is expected to achieve many improvements in the surgical waiting lists' length. More specifically, the model suggests an increase of surgical supply of about 49 000 interventions and less 19 000 patients waiting on the list at the end of one year, in comparison with the lived situation.

In the second experiment setting, scenario C, where 75.5% of patients accepted their SV when offered one, the impact was even more significant than in B. SVs' emission was of 265 and 200 of those were accepted. That resulted in 181 surgeries performed within the SVs' policy (77 more than in scenario B and 136 more than in scenario A), out of 720 total surgeries. The number of remaining patients on the list at the end of the year was 195, 13 units lower than the value obtained in scenario B and 32 less than in scenario A. Converting these values and comparing them to the surgical waiting list's actual situation, a 75.5% SVs' acceptance percentage increased the total number of treated patients in 95 000, while also lowering the number of patients still waiting 365 days after in 32 000.

The remaining four scenarios (D, E, F, G) also resulted in (more or less, depending on the case) meaningful improvements to the overall waiting list's statistics, as it can be observed in the above table. In fact, any of the considered experiences would have resulted in less patients still waiting for surgery at the end of the simulated year than the ones registered for the control setting. Scenarios C and G, however, reflect the impressive extent to which behavioural changes can indeed affect the healthcare system's functioning, as they were associated with the highest acceptance percentages in this study.

On the other hand, the SVs' acceptance percentage currently registered in Portugal is still quite far from the ones practiced in these top two scenarios. Despite that, scenarios B and E showed that even smaller increases (and perhaps more realistic for the near future) can induce valuable changes in the access to surgical care.

Therefore, and in order to achieve the improvements explored above, public health authorities should make efforts to refine the SVs' programme policies, especially aiming at minimizing the obstacles felt by patients to accept their transfer (and, consequently, at reducing their perceived barriers). Thus, offering more options of hospitals closer to patients' residential area, as well as more transportation and accommodation aids are examples of possible measures to reduce the percentage of transfer refusals. Besides perceived barriers, cues to action also proved to be an important contributing factor to patients' intention to behave. Strategies targeting this latter construct could also be effective – some examples may be: recommending family doctors (or the doctors who initially proposed surgery at the home hospital) to incentivize patients' acceptance of the SV, as well as promoting the general population's knowledge about the benefits of this programme.

Nevertheless, it is important to reinforce that the estimations presented in the above tables were extracted from a simulation model which has a certain degree of error associated to it, as it was observed with the outputs obtained for scenario A that, although close to what was experienced in reality, showed, nonetheless, slight discrepancies. Such differences emerge, naturally, from the fact that any model is a simplified representation of a given real-world problem – and the statechart represented in Figure 14 is, of course, no exception. The fact that no patients of priority level three requested an SV is an example of a limitation in the present simulation. These patients were also not considered in subpopulations *patients181*, *patients182*, and *patients183*, as their little TMRG made it less probable to have a significant number of these high priority patients transitioning from one year to the next on the waiting list. In addition, and despite it being represented in the statechart, no cancelled enrolments on the waiting list were actually readmitted during the run time of the model.

Besides that, other approximations were made that justify some of the differences observed between estimated and real values. While the absolute numbers of patients on the surgery waiting list, both entering in 2019 and transitioning from 2018, were found in national-wide reports, the proportion of patients to be distributed among the subpopulations *patients181*, *patients182*, and *patients183* was computed based on the waiting list of a particular Portuguese SNS hospital. Furthermore, it was also based on this hospital's surgery waiting list that the waiting times implemented as model transitions "Wait_P1", "Wait_P2", "Wait_P3", "Wait_P1_2", and "Wait_P2_2" were defined. It is reasonable to consider that the waiting list of one particular hospital may not completely translate the dynamics of the national list and, therefore, this approximation may account for part of the discrepancies observed. Based on the information made available by the Ministry of Health [18], it is possible to know that the average waiting time until surgery was of 3.3 months in 2019, while the median stood at 3.5 months. Moreover, 32.1% of patients had their surgeries performed after the TMRG established by law. Based on the hospital's data, the average waiting time was of 5.6 months, the median of 3 months, and 30.2% of all patients were treated after their TMRG. Such variations might have resulted in the minor errors detected in scenario A, that can also be extrapolated to the results of scenarios B to G. Another adaptation is the fact that the waiting time data used as reference concerned the list of patients still waiting for surgery at the end of 2018, and not the waiting times of a set of patients whose surgery had already been performed, which might explain the lower median time of this database when compared to national values.

The threshold for SVs' acceptances, the percentages of refusals and no answers, as well as the percentage of enrolment cancellations, were all based on national values. No information regarding the time a patient is expected to wait after accepting an SV and until the surgery is performed was found, which meant that the "Undergoes_Surgery" transition had to be defined only resorting to the feedback of a patient who had been offered an SV in the past – a countdown of 30 days was, then, defined. The lack of support regarding this timestamp is another limitation of the simulation model.

All things considered, and despite the discrepancies induced by some of these model's approximations, this research contributed with sound tools to help policymakers understanding the impact that patients' behavioural changes could have on the evolution of the Portuguese surgery waiting list over the years. Understanding the main factors responsible for patients' decision-making regarding SVs – through the HBM causal model – is, therefore, essential to provide valuable insight to health authorities regarding the type of measures that have the potential to significantly impact behaviour. That would save resources, while increasing the effectiveness of this transfer programme. The real-life effects of such policy adjustments might also be predicted and evaluated, with the help of the developed AnyLogic simulation.

According to the simulation model, once effective measures are put into practice leading to higher levels of SVs' acceptance, significant improvements in the number of patients waiting for surgery can be achieved, as explored above in this chapter. Following the Portuguese strategy implemented with SIGIC, of better distributing surgical demand among the already existing (public and private) supply capacity, this decrease in the waiting list's length should also cause a meaningful reduction in patients' waiting times. Such changes would, ultimately, result in a more accessible and responsive healthcare service, as well as in a healthier population.

Chapter 6

Conclusion

The work developed in the present dissertation contributes with two useful tools to the resolution of a concerning behavioural problem, regarding the lack of SVs' acceptance by the population of patients on the Portuguese waiting list for surgery. Although excessively long hospital waiting lists are a relatively explored matter, it is not so usual to emphasize behavioural aspects as a main concern regarding the functioning of healthcare delivery services and programmes – a gap this study tried to fill.

On the other hand, despite some BOR studies have already been applied to other healthcare contexts, such as vaccination intake, consumption of resources and services, and compliance with screening, it has not been used to study patients' decision-making regarding hospital transfers. Therefore, this dissertation emerges as an important source of knowledge regarding this subject.

The first valuable tool presented is composed by a measurement instrument of patients' a priori intention to accept an SV and its resultant causal model, where the determining factors for such intention to behave can be identified. Based on that, public health authorities and policymakers are able to create and implement more reasoned reforms or adjustments to the SVs' programme, which can achieve more effective results by directly targeting the causes of patients' intention to accept their transfer offerings.

It was the first time the HBM model was adapted to this particular health-related context and, therefore, both the questionnaire and the SEM model form a novel and important set of instruments that can be used in further studies concerning this problem, or that may even be adapted to investigate other aspects of stakeholders' decisions related to health topics.

In addition to modelling behaviour contributing factors, this dissertation also contains a simulation model that could be used to predict the impact certain behavioural changes can possibly attain. By mimicking the general waiting process, this model is not only useful to investigate variations of SVs' acceptance, as it was done in this dissertation, but it can also be used to test other policy changes regarding patients' pathways in the surgical waiting list. Therefore, it constitutes another useful tool to guide policymakers in the pursuit of a satisfactorily responsive and effective national health system, as it is stated in SIGA's objectives.

Overall, this research exposes the importance and potential of OR (its modelling and simulation techniques) to study behaviour, namely when applied to healthcare, and to help improve systems where peoples' decision-making and attitudes play a critical role in the outcomes obtained.

However, the developed research also presents its limitations. First of all, the sample analysed during the quantitative HBM-related study was considerably small. For future studies, a bigger sample would benefit the reliability of the obtained results. Moreover, other psychological models of health behaviour could be explored, since HBM does not a priori establish a mathematical formulation regarding its constructs' relationships. Another important aspect of this study that it is important to keep in mind is the fact that, in general, the participants who composed the sample were not actually enrolled on the surgery waiting list at the time of questioning. Therefore, their answers were mainly a result of what they think their feelings would be regarding a hypothetical SV offering – which may or may not coincide with what would happen in the real prospect of receiving an SV. Also because of that, the questionnaire ultimately measured people's intention to behave, rather than behaviour itself – which had to be afterwards extrapolated to be considered in the simulation.

An improvement to this study would have been to consider a sample of patients currently enrolled for surgery, who were eligible to receive a transfer offer in reality, and to monitor their waiting process in order to observe their true decisions regarding SVs. In that case, not only the obtained answers would be more accurate and reliable, but it would also be possible to directly measure behaviour.

Furthermore, some approximations made during the implementation of the simulation model could be perfected. For example, using waiting times from already treated patients at a national level, instead of data regarding waiting patients from a single Portuguese hospital (which may not entirely match the country-wide waiting trend) would be beneficial. It would also be important to have access to information regarding the waiting times until surgery for patients who already accepted their SVs, which was not possible to find for this research.

Moreover, the accuracy of the simulation model's outcomes could have been affected by other simplifications adopted. For instance, input values had to be approximated so that the used software could process them. Besides that, the TMRGs considered for each level of clinical priority only referred to cases of "general pathologies", and not to cardiac or oncologic diseases (which have specific maximum waiting times, as shown in Table 1). Patients with a clinical priority of level three were only considered in the subpopulation *patients19*, and not for patients transitioning from 2018. Other mechanisms of the waiting list, such as patients' request for a third SV after two previous refusals and their re-enrolment after a non-existing answer were also not considered during the simulation. The population of agents was created based on a study from 2007 and, because of that, may not completely correspond to the current waiting list's demographic description.

Nevertheless, it is part of the modelling process to simplify some aspects of the problem being tackled. Such simplifications were considered to be acceptable, since they did not compromise the model's objective of representing the surgery waiting list's general functioning. Therefore, the model fulfilled its purpose of proving the impact that patients' behaviour changes regarding SVs could have on the waiting list's dynamics.

Although this particular research embraced the assumption that the Portuguese healthcare system would have enough capacity to deliver the total number of surgeries obtained for the multiple considered scenarios, future studies can also include resource restrictions in order to evaluate if services would, indeed, be able to respond properly.

Bibliography

[1] J. de Almeida Simões, G. F. Augusto, and I. Fronteira, *Portugal: Health system review*. Health Systems in Transition, 2017, vol. 19, no. 2.

[2] P. P. Barros, S. R. Machado, and J. de Almeida Simões, *Portugal: Health system review*. Health Systems in Transition, 2011, vol. 13, no. 4.

[3] PORDATA, *Taxa bruta de mortalidade e taxa de mortalidade infantil*, (Accessed last: 13.03.2021). [Online]. Available: https://www.pordata.pt/Portugal/Taxa+bruta+de+mortalidade+e+taxa+de+mortalidade+infantil-528

[4] *Manual de Gestão de Inscritos para Cirurgia*. UCGIC, 2011, vol. 1 - Princípios Gerais.

[5] *Auditoria ao Acesso a Cuidados de Saúde no Serviço Nacional de Saúde*. Tribunal de Contas, 2017.

[6] J. Hurst and L. Siciliani, "Tackling excessive waiting times for elective surgery: a comparison of policies in twelve oecd countries," 2003.

[7] P. P. Barros, "As listas de espera para intervenção cirúrgica em portugal," in *iprisverbis*, no. 4. Instituto Português de Relações Internacionais e Segurança, 2008.

[8] OECD, *Waiting Times for Health Services*, 2020. [Online]. Available: https://www.oecd-ilibrary.org/content/publication/242e3c8c-en

[9] *Auditoria ao Programa Especial de Combate às Listas de Espera Cirúrgicas -PECLEC*. Tribunal de Contas, 2005.

[10] *Resolução do Conselho de Ministros n.º 79/2004*. Diário da República n.º 147/2004, Série I-B de 2004-06-24. [Online]. Available: https://data.dre.pt/eli/resolconsmin/79/2004/06/24/p/dre/pt/html

[11] *SIGA SNS – Sistema Integrado de Gestão do Acesso no Serviço Nacional de Saúde*, (Accessed last: 13.03.2021). [Online]. Available: https://www.sns.gov.pt/sns/sistema-integrado-de-gestao-do-acesso-no-servico-nacional-de-saude-siga-sns/

[12] *Portaria n.º 1529/2008*. Diário da República n.º 249/2008, Série I de 2008-12-26. [Online]. Available: https://data.dre.pt/eli/port/1529/2008/12/26/p/dre/pt/html

[13] *Portaria n.º 153/2017*. Diário da República n.º 86/2017, Série I de 2017-05-04. [Online]. Available: https://data.dre.pt/eli/port/153/2017/05/04/p/dre/pt/html

[14] *Portaria n.º 45/2008*. Diário da República n.º 10/2008, Série I de 2008-01-15. [Online]. Available: https://data.dre.pt/eli/port/45/2008/01/15/p/dre/pt/html

[15] *Portaria n.º 147/2017*. Diário da República n.º 82/2017, Série I de 2017-04-27. [Online]. Available: https://data.dre.pt/eli/port/147/2017/04/27/p/dre/pt/html

[16] *Medical Dictionary for the Health Professions and Nursing*, 2012, (Accessed last: 13.03.2021). [Online]. Available: https://medical-dictionary.thefreedictionary.com/patient+episode

[17] *Relatório síntese da atividade cirúrgica programada*. UCGIC, 2015-2012.

[18] *Relatório Anual - Acesso a Cuidados de Saúde nos Estabelecimentos do SNS e entidades*. Ministério da Saúde, 2019.

[19] A. Ferreira, C. Antunes, E. Pires, M. L. M. Viegas, and P. Gomes, *Estudo dos motivos de recusa de transferência - Inquérito Telefónico*. UCGIC, 2008.

[20] *OR and EURO*. The Association of European Operational Research Societies, (Accessed last: 13.03.2021). [Online]. Available: https://www.euro-online.org/web/pages/301/or-and-euro

[21] T. Monks, "Operational research as implementation science: definitions, challenges and research priorities," *Implementation Science*, vol. 11, no. 1, pp. 1–10, 2015.

[22] A. Rais and A. Viana, "Operations research in healthcare: a survey," *International transactions in operational research*, vol. 18, no. 1, pp. 1–31, 2011.

[23] M. Kunc, P. Harper, and K. Katsikopoulos, "A review of implementation of behavioural aspects in the application of or in healthcare," *Journal of the Operational Research Society*, vol. 71, no. 7, pp. 1055–1072, 2020.

[24] L. A. Franco and R. P. Hämäläinen, "Engaging with behavioral operational research: On methods, actors and praxis," in *Behavioral Operational Research*. Springer, 2016, pp. 3–25.

[25] R. P. Hämäläinen, J. Luoma, and E. Saarinen, "On the importance of behavioral operational research: The case of understanding and communicating about dynamic systems," *European Journal of Operational Research*, vol. 228, no. 3, pp. 623–634, 2013.

[26] S. Brailsford and J. Vissers, "Or in healthcare: A european perspective," *European journal of operational research*, vol. 212, no. 2, pp. 223–234, 2011.

[27] S. Brailsford and B. Schmidt, "Towards incorporating human behaviour in models of health care systems: An approach using discrete event simulation," *European journal of operational research*, vol. 150, no. 1, pp. 19–31, 2003.

[28] A. Fügener, S. Schiffels, and R. Kolisch, "Overutilization and underutilization of operating rooms-insights from behavioral health care operations management," *Health care management science*, vol. 20, no. 1, pp. 115–128, 2017.

[29] R. Sachdeva, T. Williams, and J. Quigley, "Mixing methodologies to enhance the implementation of healthcare operational research," *Journal of the Operational Research Society*, vol. 58, no. 2, pp. 159–167, 2007.

[30] G. Royston, A. Dost, J. Townshend, and H. Turner, "Using system dynamics to help develop and implement policies and programmes in health care in england," *System Dynamics Review: The Journal of the System Dynamics Society*, vol. 15, no. 3, pp. 293–313, 1999.

[31] M. Conner and P. Norman, "Health behaviour: Current issues and challenges," 2017.

[32] M. Conner, "Health behaviors," *International Encyclopedia of the Social & Behavioral Sciences*, 12 2015.

[33] S. Mkhitaryan, P. J. Giabbanelli, N. K. de Vries, and R. Crutzen, "Dealing with complexity: How to use a hybrid approach to incorporate complexity in health behavior interventions," *Intelligence-Based Medicine*, vol. 3, p. 100008, 2020.

[34] B. Australia, B. Wright, and P. Bragge, "Stakeholder consultation to improve behaviour change."

[35] M. Pearson, T. Monks, A. Gibson, M. Allen, A. Komashie, A. Fordyce, F. Harris-Golesworthy, M. Pitt, S. Brailsford, and K. Stein, "Involving patients and the public in healthcare operational research—the challenges and opportunities," *Operations Research for Health Care*, vol. 2, no. 4, pp. 86–89, 2013.

[36] R. Sachdeva, T. Williams, and J. Quigley, "Mixing methodologies to enhance the implementation of healthcare operational research," *Journal of the Operational Research Society*, vol. 58, no. 2, pp. 159–167, 2007.

[37] P.-J. Hsieh and H.-M. Lai, "Exploring peoples intentions to use the health passbook in self-management: An extension of the technology acceptance and health behavior theoretical perspectives in health literacy," *Technological Forecasting and Social Change*, vol. 161, p. 120328, 2020.

[38] S. C. Brailsford, "Healthcare: Human behavior in simulation models," in *Behavioral operational research*. Springer, 2016, pp. 263–280.

[39] F. Ashoori, M. Karimi, H. Mokarami, and M. Seif, "Using health belief model to predict oral health behaviors in girl students: A structural equation modeling," *Pediatric Dental Journal*, vol. 30, no. 1, pp. 24–32, 2020.

[40] S. C. Brailsford, P. R. Harper, and J. Sykes, "Incorporating human behaviour in simulation models of screening for breast cancer," *European Journal of Operational Research*, vol. 219, no. 3, pp. 491–507, 2012.

[41] S. C. Brailsford, J. Sykes, and P. R. Harper, “Incorporating human behavior in healthcare simulation models,” in *Proceedings of the 2006 Winter Simulation Conference*. IEEE, 2006, pp. 466–472.

[42] A. P. G. Scheidegger, T. F. Pereira, M. L. M. de Oliveira, A. Banerjee, and J. A. B. Montevechi, “An introductory guide for hybrid simulation modelers on the primary simulation methods in industrial engineering identified through a systematic review of the literature,” *Computers & Industrial Engineering*, vol. 124, pp. 474–492, 2018.

[43] S. C. Brailsford, “Modeling human behavior-an (id) entity crisis?” in *Proceedings of the Winter Simulation Conference 2014*. IEEE, 2014, pp. 1539–1548.

[44] M. A. Lopes, Á. S. Almeida, and B. Almada-Lobo, “Forecasting the medical workforce: a stochastic agent-based simulation approach,” *Health care management science*, vol. 21, no. 1, pp. 52–75, 2018.

[45] R. Lebcir, E. Demir, R. Ahmad, C. Vasilakis, and D. Southern, “A discrete event simulation model to evaluate the use of community services in the treatment of patients with parkinson’s disease in the united kingdom,” *BMC health services research*, vol. 17, no. 1, pp. 1–14, 2017.

[46] T. E. Day, N. Ravi, H. Xian, and A. Brugh, “Sensitivity of diabetic retinopathy associated vision loss to screening interval in an agent-based/discrete event simulation model,” *Computers in biology and medicine*, vol. 47, pp. 7–12, 2014.

[47] J. Viana, S. C. Brailsford, V. Harindra, and P. R. Harper, “Combining discrete-event simulation and system dynamics in a healthcare setting: A composite model for chlamydia infection,” *European Journal of Operational Research*, vol. 237, no. 1, pp. 196–206, 2014.

[48] L. Alhalaseh, H. Fayoumi, and B. Khalil, “The health belief model in predicting healthcare workers’ intention for influenza vaccine uptake in jordan,” *Vaccine*, vol. 38, no. 46, pp. 7372–7378, 2020.

[49] M. C. Wong, E. L. Wong, J. Huang, A. W. Cheung, K. Law, M. K. Chong, R. W. Ng, C. K. Lai, S. S. Boon, J. T. Lau *et al.*, “Acceptance of the covid-19 vaccine based on the health belief model: A population-based survey in hong kong,” *Vaccine*, vol. 39, no. 7, pp. 1148–1156, 2021.

[50] E. Knekta, C. Runyon, and S. Eddy, “One size doesn’t fit all: Using factor analysis to gather validity evidence when using surveys in your research,” *CBE—Life Sciences Education*, vol. 18, no. 1, p. rm1, 2019.

[51] D. D. Suhr, “Exploratory or confirmatory factor analysis?” 2006.

[52] R. D. Bock, *Multivariate statistical methods in behavioral research*. Scientific Software International, 1985.

[53] M. W. Watkins, “Exploratory factor analysis: A guide to best practice,” *Journal of Black Psychology*, vol. 44, no. 3, pp. 219–246, 2018.

[54] E. T. Pribadi and S. R. Devy, “Application of the health belief model on the intention to stop smoking behavior among young adult women,” *Journal of Public Health Research*, vol. 9, no. 2, 2020.

[55] H. Le, T. Dang, R. Ware, D. Phung, P. Thai, P. Sly, and P. Le An, “Using the health beliefs model to explore children’s attitudes and beliefs on air pollution,” *Public Health*, vol. 196, pp. 4–9, 2021.

[56] J. Marôco, *Análise de equações estruturais: Fundamentos teóricos, software & aplicações*. ReportNumber, Lda, 2010.

[57] N. K. Bowen and S. Guo, *Structural equation modeling*. Oxford University Press, 2011.

[58] M. de Fátima Salgueiro, *Modelos com Equaçoes Estruturais*. Sociedade Portuguesa de Estatística, 2012.

[59] D. J. Mundfrom, D. G. Shaw, and T. L. Ke, “Minimum sample size recommendations for conducting factor analyses,” *International Journal of Testing*, vol. 5, no. 2, pp. 159–168, 2005.

[60] X. Qiao, L. Ji, Y. Jin, H. Si, Y. Bian, W. Wang, and C. Wang, “Development and validation of an instrument to measure beliefs in physical activity among (pre) frail older adults: an integration of the health belief model and the theory of planned behavior,” *Patient Education and Counseling*, 2021.

[61] K. Hae-Young, “Statistical notes for clinical researchers: Chi-squared test and fisher’s exact test,” *rde*, vol. 42, no. 2, pp. 152–155, 2017. [Online]. Available: http://www.e-sciencecentral.org/articles/?scid=1090217

[62] P. E. McKnight and J. Najab, “Mann-whitney u test,” *The Corsini encyclopedia of psychology*, pp. 1–1, 2010.

[63] D. Child, *The Essentials of Factor Analysis*, 2nd ed., L. C. E. Ltd, Ed., 1990.

[64] R. L. Gorsuch, *Factor Analysis*, 2nd ed., G. S. of Psychology, Ed. Lawrence Erlbaum Associates, Inc, 1983.

[65] H. F. Kaiser, “An index of factorial simplicity,” *Psychometrika*, vol. 39, no. 1, pp. 31–36, 1974.

[66] M. Courtney and R. Gordon, “Determining the number of factors to retain in efa: using the spss r-menu v2 0 to make more judicious estimations,” *Practical Assessment, Research, and Evaluation*, vol. 18, no. 1, p. 8, 2013.

[67] J. F. Hair, “Multivariate data analysis,” 2009.

[68] T. A. Brown, *Confirmatory factor analysis for applied research*. Guilford publications, 2015.

[69] D. George and P. Mallery, *IBM SPSS statistics 26 step by step: A simple guide and reference.* Routledge, 2019.

[70] F. Yang-Wallentin, K. G. Jöreskog, and H. Luo, “Confirmatory factor analysis of ordinal variables with misspecified models,” *Structural Equation Modeling*, vol. 17, no. 3, pp. 392–423, 2010.

[71] C.-H. Li, “The performance of ml, dwls, and uls estimation with robust corrections in structural equation models with ordinal variables.” *Psychological methods*, vol. 21, no. 3, p. 369, 2016.

[72] J. T. Newsom, “Structural equation modeling with mplus,” *Early Head Start*, 2004.

[73] I. Grigoryev, *AnyLogic 8 in Three Days*, 2021.

Appendix A

HBM Questionnaire

This appendix presents the HBM-based questionnaire developed to measure patients' willingness to accept an SV. This survey was developed in Google Forms and it was shared on social media (Facebook and LinkedIn), as well as door-to-door.

The questionnaire was originally written in Portuguese. However, below it is presented a translated version to English.

A.1 Study of the determining factors for SVs' acceptance

My name is Madalena Almeida and I am currently developing my master's thesis in Biomedical Engineering, at Instituto Superior Técnico. The objective of this research is to study the system of surgical vouchers implemented in Portugal. Despite offering faster treatment for patients on the waiting list, these vouchers are often not used by patients.

One of the problems faced by SNS patients is a long waiting time for elective surgery. Currently, if a hospital is unable to provide surgical care within the time period established by law - the so-called Maximum Guaranteed Response Times (TMRGs) -, a surgery voucher is issued and sent to the patient of concern, so that he/she can be treated at another facility, with lower response times. A list of hospitals (public and/or private) is also sent to patients, for them to choose where to "spend" the voucher and schedule their surgery (with no additional costs). If patients do not accept the voucher, they remain on their home hospital's waiting list, with no guarantee of having their surgery scheduled in the near future or within the TMRG.

The following questions will be used as a basis for the development of an individual health behaviour model, in this case, regarding surgical vouchers' acceptance.

Collected responses are anonymous and will only be used within the scope of this work.

A.1.1 Demographic characterization and past experience on the surgery waiting list

- Q01: "Age?" ☐ 0-20 ☐ 21-40 ☐ 41-60 ☐ 61-80 ☐ Above 80
- Q02: "Sex?" ☐ Female ☐ Male
- Q03: "District?" ☐ Aveiro ☐ Beja ☐ Braga ☐ Bragança ☐ Castelo Branco ☐ Coimbra ☐ Évora ☐ Faro ☐ Guarda ☐ Leiria ☐ Lisboa ☐ Portalegre ☐ Porto ☐ Santarém ☐ Setúbal ☐ Viana do Castelo ☐ Viseu ☐ Região Autónoma dos Açores ☐ Região Autónoma da Madeira
- Q04: "Education level?" ☐ None ☐ Primary school ☐ Middle school ☐ High school ☐ University education
- Q05: "In general, how do you evaluate your health status?"
  ☐ 1 (very bad) ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 (excellent)
- Q06: "Have you ever been enrolled on the surgery waiting list before?" ☐ Yes ☐ No
- Q07: "If yes, have you ever received an SV? Did you use it?"
  ☐ I received it and used it ☐ I received it, but did not use it ☐ I did not receive it

A.1.2 HBM-related questions

If you have ever received an SV in the past, you should answer the following questions according to what you recall from that experience. If you have never been enrolled on the waiting list for elective surgery nor received one of these vouchers, imagine yourself in that situation and answer the following questions in accordance with that hypothetical scenario.

Please choose your level of agreement with the following statements: where: 1 = completely disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = completely agree.

- Q1: "It is likely that my health status significantly deteriorates if my surgery is not performed within the TMRG provided by law."
  ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
- Q2: "A long waiting time may significantly compromise my post-operative outcomes and my recovery."
  ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
- Q3: "Health complications resulting from a long waiting time can severely affect my ability to work."
  ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
- Q4: "Health complications resulting from a long waiting time can severely affect my ability to perform common daily tasks."
  ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q5: ”Health complications resulting from a long waiting time can severely affect my family life.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q6: ”Health complications resulting from a long waiting time can severely affect my quality of life.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q7: ”Accepting the surgery voucher means that my health problem will be solved faster.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q8: ”Accepting the surgery voucher will enable me to regain my previous quality of life more quickly.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q9: ”Accepting the surgery voucher will decrease the likelihood of deterioration of my health status.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q10: ”I would not like to be treated by a medical team other than the one from my home hospital.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q11: ”I would not like to be treated in a hospital other than my home hospital.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q12: ”Going to a hospital outside my residence area is too much of an inconvenience for me.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q13: ”Having surgery in a hospital outside my residence area is too much of an inconvenience for me.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q14: ”My work and/or family situation would cause me to postpone my surgery as much as possible.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q15: ”The anxiety resulting from the surgery date’s approaching would cause me to postpone the procedure as much as possible.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q16: ”Family members’ support to accept the surgery voucher would influence my decision.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q17: ”Friends’ support to accept the surgery voucher would influence my decision.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q18: ”The support of my home hospital’s medical team or my family doctor to accept the voucher would influence my decision.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q19: ”Contacting with other patients who had previously scheduled their surgeries though the surgery vouchers’ system would influence my decision.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q20: ”A deterioration of my health condition during the waiting period would influence my decision.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q21: ”Upon receiving a surgery voucher, I would be able to contact and schedule the surgery myself at one of the hospitals from the options list.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q22: ”If it was not clear to me how I should proceed to use the voucher, I would be able to search information regarding that process.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q23: ”I try to maintain a healthy lifestyle.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q24: ”I try to maintain a regular physical activity.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q25: ”I try to maintain a healthy diet.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q26: ”I regularly search for information to help me stay healthy.”

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

– Q27: ”If you would receive a surgery voucher in the future, would you be willing to use it?”

☐ 1 (definitely no) ☐ 2 ☐ 3 ☐ 4 ☐ 5 (definitely yes)

Thank you for your participation!

Appendix B

Simulation Model

This appendix contains the AnyLogic statecharts for agents belonging to subpopulations *patients181*, *patients182*, and *patients183*.

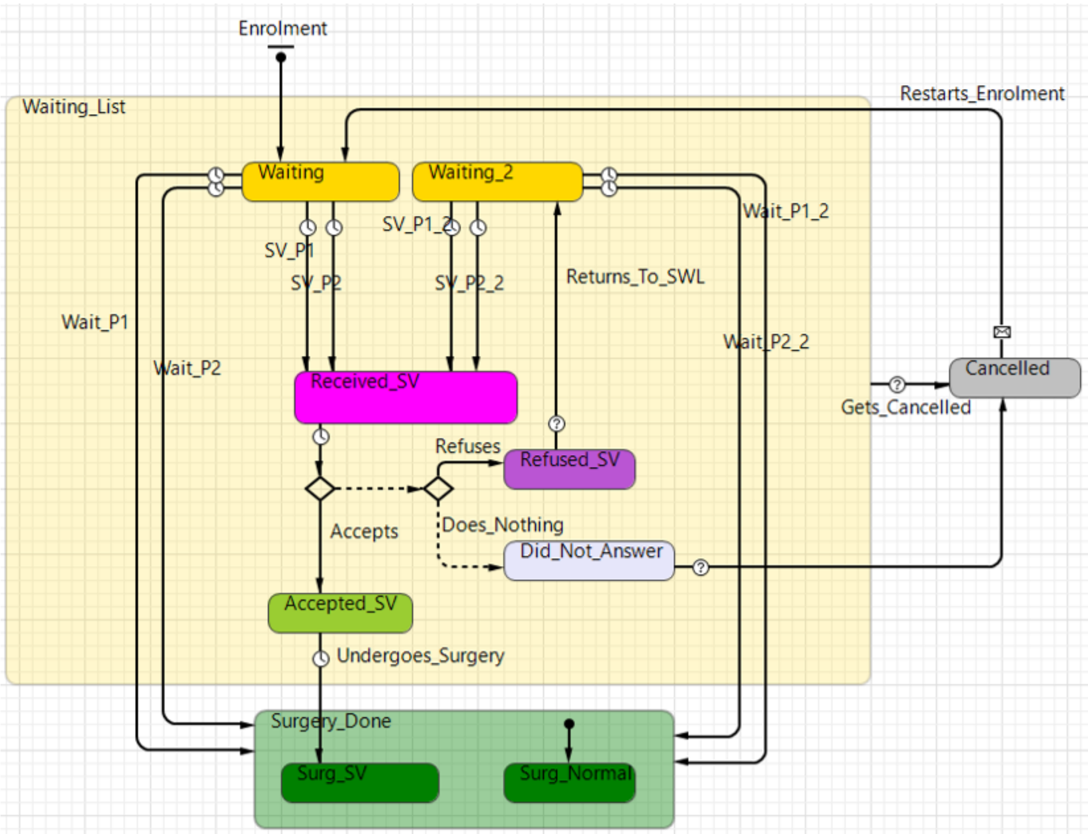

Figure 21: Statechart developed in AnyLogic to define *Patient181* agents' states and transitions in the simulation, as a representation of patients' possible pathways in the Portuguese surgery waiting list.

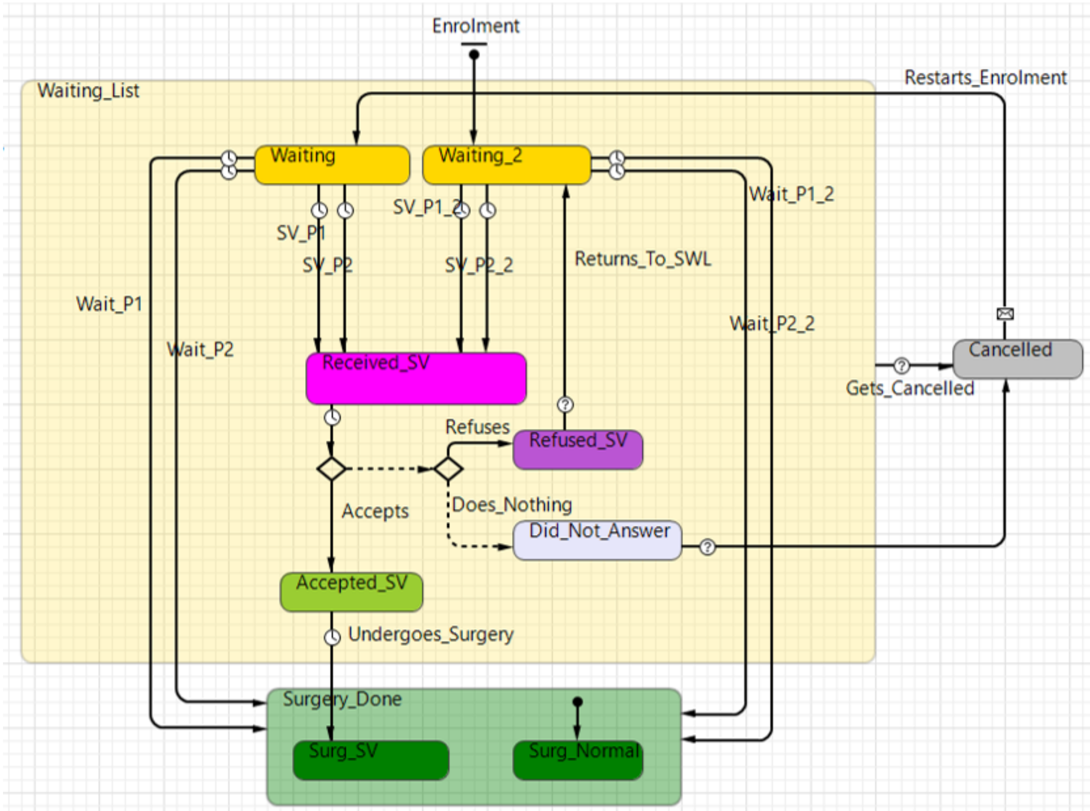

Figure 22: Statechart developed in AnyLogic to define *Patient182* agents' states and transitions in the simulation, as a representation of patients' possible pathways in the Portuguese surgery waiting list.

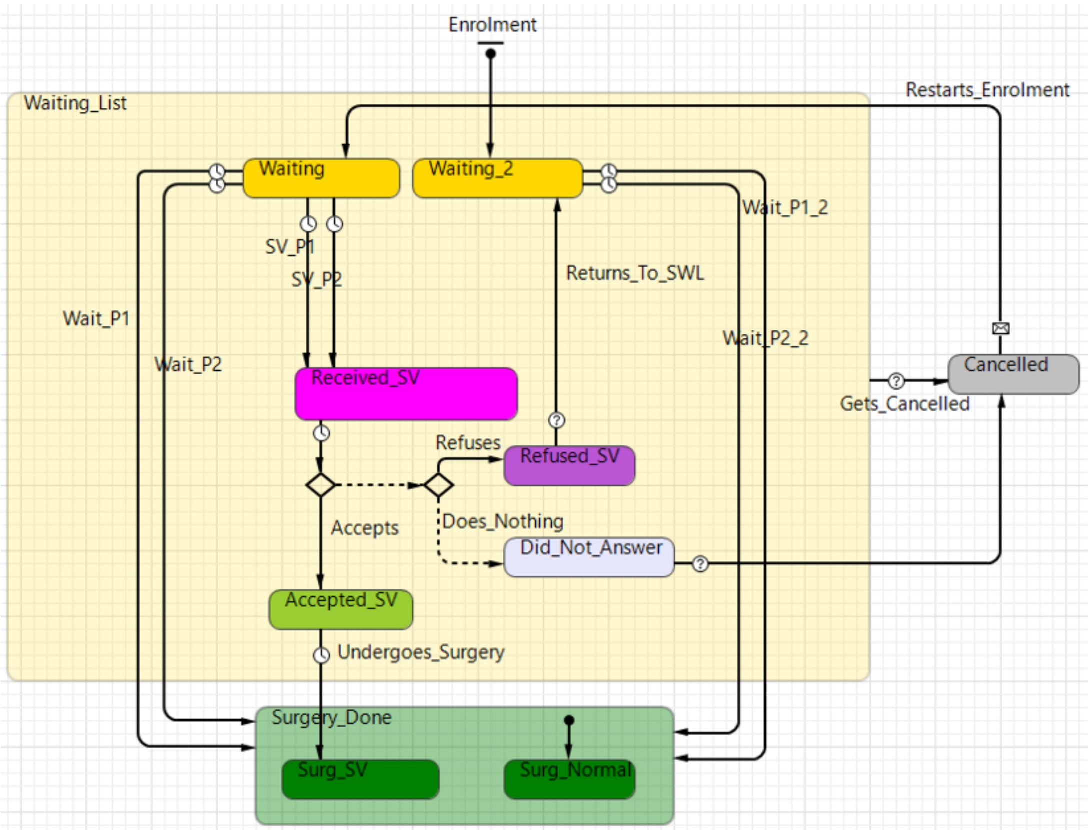

Figure 23: Statechart developed in AnyLogic to define *Patient183* agents' states and transitions in the simulation, as a representation of patients' possible pathways in the Portuguese surgery waiting list.

Appendix C

Quantitative Study

This appendix comprises additional results of the HBM-based questionnaire's EFA.

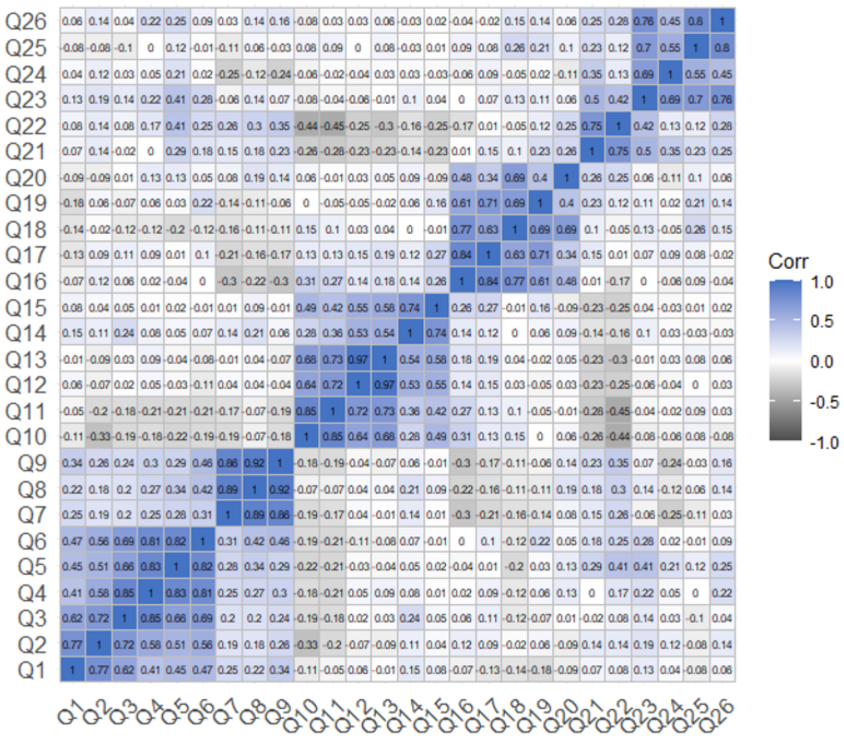

Figure 24: Matrix of the polychoric correlation coefficients among variables Q1 to Q26, corresponding to those same questionnaire items, obtained during EFA.

Table 18: Observed variables (excluding Q21 and Q22), their corresponding factor loading in all five considered factors (PA1 to PA5), as well as their communality and uniqueness values, obtained during EFA. Factor loading coefficients higher than the 0.50 cut-off are written in bold.

Variable	Factor Loading Coefficients (λ_{ij})					Communality (h^2)	Uniqueness (u^2)
	PA1	PA2	PA3	PA4	PA5		
Q1	0.04	**0.55**	-0.13	-0.07	0.01	0.33	0.67
Q2	-0.04	**0.71**	0.01	-0.03	-0.09	0.48	0.52
Q3	0.02	**0.88**	-0.01	-0.09	-0.07	0.74	0.26
Q4	0.00	**0.77**	0.06	0.06	0.09	0.64	0.36
Q5	-0.04	**0.67**	-0.02	0.22	0.12	0.58	0.42
Q6	-0.08	**0.64**	0.08	0.05	0.20	0.53	0.47
Q7	0.01	0.01	-0.07	-0.10	**0.85**	0.76	0.24
Q8	0.07	-0.01	-0.01	0.06	**0.95**	0.90	0.10
Q9	-0.07	0.06	0.00	-0.02	**0.86**	0.78	0.22
Q10	**0.73**	-0.16	0.08	-0.04	-0.05	0.61	0.39
Q11	**0.77**	-0.14	0.03	-0.01	-0.06	0.65	0.35
Q12	**0.90**	0.04	-0.05	-0.02	0.03	0.80	0.20
Q13	**0.92**	0.02	-0.02	0.03	0.01	0.84	0.16
Q14	**0.57**	0.16	0.02	0.05	0.15	0.38	0.62
Q15	**0.60**	0.13	0.08	0.07	0.05	0.39	0.61
Q16	0.13	0.10	**0.85**	-0.09	-0.14	0.83	0.17
Q17	0.08	0.13	**0.76**	-0.02	-0.07	0.63	0.37
Q18	-0.06	-0.16	**0.84**	0.07	0.06	0.72	0.28
Q19	-0.12	0.01	**0.77**	0.05	0.01	0.58	0.42
Q20	-0.01	-0.07	**0.55**	0.01	0.20	0.30	0.70
Q23	0.00	0.11	-0.02	**0.84**	-0.02	0.73	0.27
Q24	0.02	0.09	-0.12	**0.61**	-0.26	0.43	0.57
Q25	0.02	-0.14	0.09	**0.78**	0.00	0.63	0.37
Q26	0.01	-0.02	0.01	**0.78**	0.08	0.61	0.39