Optimizing Flow in the Healthcare System for Older Adults

Arntzen, Rebekka Jantje

2025

DOI (link to publisher)
10.5463/thesis.1002

document version
Publisher's PDF, also known as Version of record

Optimizing Flow in the Healthcare System for Older Adults

ISBN: 978-94-6473-689-2

Typeset by LaTeX.
Printed by: Ipskamp Printing
Cover design by: Persoonlijk Proefschrift

VRIJE UNIVERSITEIT

Optimizing Flow in the Healthcare System for Older Adults

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Bètawetenschappen
op woensdag 12 februari 2025 om 13.45 uur
in een bijeenkomst van de universiteit,
De Boelelaan 1105

door

Rebekka Jantje Arntzen

geboren te Amsterdam

promotoren:	prof.dr. R.D. van der Mei prof.dr. S. Bhulai
copromotoren:	dr. R. Bekker prof.dr. B.M. Buurman-van Es
promotiecommissie:	prof.dr. S. Brailsford prof.dr.ir. E.W. Hans dr. C.J. Jagtenberg prof.dr. G.M. Koole prof.dr. K.D.P.W.B. Nanayakkara

The ultimate moral test of any government is the way it treats three groups of its citizens. First, those in the dawn of life — our children. Second, those in the shadows of life — our needy, our sick, our disabled. Third, those in the twilight of life — our aged.

Hubert Humphrey

Voorwoord

Mijn promotietraject startte op 1 mei 2020 vanaf een bureau op mijn slaapkamer, al ruim een maand vol in lockdown. Die dag begon met een Zoom-meeting met Rob die een halfuur duurde, waarna ik een beetje beduusd om me heen keek: mijn promotieonderzoek was begonnen! Inmiddels zijn we meer dan vier jaar verder en is mijn proefschrift voltooid. Het is een mooie bundeling van onderzoeken geworden waar ik trots op ben. Bij mijn onderzoeken heb ik de zorgpraktijk als uitgangspunt genomen, waardoor de meeste conclusies direct richting geven aan beleid. Er zijn zoveel mogelijkheden om de doorstroom in de ouderenzorg te verbeteren, dat ik er vertrouwen in heb dat deze kansen ook zullen worden benut.

Het tot stand komen van dit proefschrift is ook de verdienste van vele anderen. Ten eerste wil ik graag mijn dank uitspreken aan R^3: Rob van der Mei, René Bekker en Tim R. de Boer. Op donderdagochtenden vormden wij het R^4 overleg. Rob, graag wil ik jou bedanken voor je vertrouwen in mij, je humor en positiviteit. Bij ontelbaar veel projectmeetings heb ik gezien hoe jij anderen enthousiast krijgt voor ons wiskundige onderzoek - een echte vakman! René, bedankt voor je betrouwbaarheid, het delen van je onuitputtelijke kennis over wachtrijtheorie en je vermogen om (wiskundige) problemen op te lossen. Inmiddels heb ik een automatisch zinnetje in mijn hoofd 'Ik kan het natuurlijk altijd nog aan René vragen!' Tim, jou wil ik speciaal bedanken voor je mentale support en je begrip. Maar zeker ook zoveel voor je gezelligheid en het fijne samenwerken. Jij was mijn maatje in het project en dat heeft ook altijd zo gevoeld!

Verder wil ik Judith bedanken voor de goede samenwerking die wij hebben gehad. Wij vulden elkaar perfect aan met onze achtergronden in geriatrie en wiskunde, waarmee we twee mooie artikelen hebben kunnen schrijven. Ook het congres in Antwerpen waar wij samen waren was voor mij een van de hoogtepunten van mijn promotietijd.

Graag wil ik ook mijn dank uitspreken aan iedereen van het DOLCE VITA project. Sandjai, Bianca, Hanna, Oscar, Robert, Ismail en Eline, bedankt voor jullie nuttige input tijdens de meetings. Hierdoor kon ik mijn ideeën en resultaten toetsen en snel schakelen. Daarnaast wil ik ook de gebruikersgroep bedanken voor hun betrokkenheid bij ons onderzoek en waardevolle inbreng.

De Stochasticsgroep van het Centrum Wiskunde & Informatica heb ik ervaren als een team van hele fijne mensen. Ik kijk warm terug op het tafeltennissen, de borrels op het dakterras en de Polder en de gesprekken bij de koffieautomaat die steevast uitliepen op een maatschappelijk debat. Met Etienne en Jan is hier voor mij een goede vriendschap ontstaan. Etienne, dat wij allebei van kletsen houden is denk ik vrij weinig mensen op de afdeling ontgaan. Bedankt voor je gezelligheid en het zijn van een stabiele factor in het roerige PhD-leven. Jan, wij zitten op precies dezelfde golflengte en ik kan ontzettend met jou lachen. Ik ben heel blij dat we nu opnieuw collega's zijn!

Daarnaast heb ik ook de A&O-groep onder leiding van Ger Koole erg gewaardeerd. De seminars hebben mijn kennis van Operations Research verdiept en mijn kijk op Machine Learning verbreed. Ook de uitjes hebben mijn onderzoekstraject leuker gemaakt, met als letterlijk hoogtepunt de jaarlijkse huttentocht naar de Alpen of Dolomieten. Aan het bezoeken van congressen zoals INFORMS (Healthcare), de Workshop on Queueing Theory en StochMod met A&O-collega's heb ik positieve herinneringen. Graag wil ik ieders bijdrage hieraan bedanken.

Furthermore, I want to express my gratitude to the doctorate committee for reading my thesis and I am looking forward to fruitful discussions during the defense.

Met een aan zekerheid grenzende waarschijnlijkheid was dit proefschrift er niet geweest zonder Thije van Barneveld. Thije, bedankt voor je suggestie om een proefschrift te gaan schrijven. Jouw vertrouwen en verwijzing naar Rob hebben mij het duwtje in de rug gegeven om het avontuur aan te gaan.

Tijdens mijn proefschriftperiode heb ik veel plezier beleefd aan ontspanning met vrienden. Yentl, mijn roeimaatje, oh nee, bootcampmaatje! Jij hebt als geen ander de rollercoaster meegekregen die ik de afgelopen jaren heb beleefd. Bedankt voor alle gezelligheid en support! Petra, Lineke, Tess, Rikst, Janneke, Vivienne, Rikst en Manon, bedankt voor de feestjes en etentjes met gezellige flessen rosé. Evelien, Babette, Anne en Marjolein, bedankt voor de knutselavonden, weekendjes weg met hottubs, en het delen van onze voorliefde voor Seanna Paul. Tessel, Mathijs en Bart, voor de etentjes en spelletjesavonden. Nienke, voor de geweldige Alpenavonturen die we samen hebben beleefd, en Romaaike, voor het boulderen en vooral voor de pizza erna. Loïs, Lenneke, Gijs, Femke, Frank en Melvin, voor het padellen en de legendarische karaokeavond!

Speciale dank wil ik uitspreken aan mijn paranimfen: de toppers Hanne en Nina. Hanne, van Halloween op onze woonboot tot après-ski op wintersport, wat hebben wij eigenlijk niet meegemaakt samen? Ik waardeer jouw pit, eerlijkheid en humor, en dat je altijd voor me klaarstaat. Nina, een recent hoogtepunt was onze roadtrip langs de oostkust van de VS. Maar het net iets minder recente barvrouw zijn op Texel zal ik ook nooit vergeten. Jij bent altijd betrokken en hebt een hart van goud!

Tot slot wil ik graag mijn familie bedanken voor hun onvoorwaardelijke steun tijdens deze voor mij tumultueuze jaren. Papa, dank je voor je interesse, je rotsvaste vertrouwen in mij en dat ik altijd bij je terechtkan. Mama, bedankt voor je warmte, gezelligheid en dat je mij motiveert om mijn hart te volgen. Siebe, ik ben zo ontzettend trots op jou! Jij hebt mij laten inzien hoe belangrijk goede zorg is en ik ben blij om met dit proefschrift hier mijn steentje aan bij te mogen dragen.

Contents

List of Abbreviations

ADL	Activities of Daily Living
ALF	Assisted Living Facilities
AM	Allocation Model
AT	Admission Turns
BIP	Binary Integer Program
CICN	Center for Indication of Care Needs
CP	Current Policy
CROSS	Checklist for Reporting of Survey Studies
DES	Discrete Event Simulation
DHA	Dutch Healthcare Authority
DTC	Diagnostics-Treatment Combination
ED	Emergency Department
FCFS	First-Come-First-Served
FP	Fast Placement
GP	General Practitioner

GR Geriatric Rehabilitation

HC Home Care

HIA Health Insurance Act

HOME No Care at Home

HOS Hospitalization

HSR Heterogeneous Service Requirements

HT Higher Tariff

IC Intermediate Care

IQP Integer Quadratic Program

KBR Knapsack-Based Routing

LoS Length-of-Stay

LTC Long-Term Care

LTCA Long-Term Care Act

MDP Markov Decision Process

MIQP Mixed Integer Quadratic Program

NH Nursing Home

NHC Nursing Home Care

NHH Nursing Home at Home

OR Operations Research

PC Personal Care

PG Psychogeriatric

PM Process Mining

PP Preferred Placement

RBB Reduced Bed-Blocking

RCM Resource Collaboration and Multitasking

SBR Skill-Based Routing

SN Statistics Netherlands

SOM Somatic

SSA Social Support Act

STC Short-Term Care

STRC Short-Term Residential Care

STRC-pall Palliative Short-Term Residential Care

STROBE STrengthening the Reporting of OBservational studies in Epidemiology

TW Triage Ward

1 Introduction

The global population is aging rapidly. The proportion of individuals aged 65 and older worldwide increased from 7 percent in 2000 to 10 percent in 2022 and is projected to reach 16 percent by 2050 [169]. In countries such as Italy, Portugal, Greece, Japan, and Korea, the share of this age group is expected to exceed one-third by 2050 [134]. This demographic shift results from declining fertility rates and improved mortality rates due to advancements in healthcare, particularly among older adults [169]. As the number of people living beyond 65 grows, ensuring adequate health and care for older adults becomes increasingly crucial.

A similar demographic transformation is occurring in the Netherlands. The ratio of individuals aged 65 and older to those aged 20 to 65 (the working-age population), known as *grey pressure*, is illustrated in Figure 1.1. This ratio has risen to approximately 34 percent and is anticipated to reach 45 percent by 2040. This shift underscores the growing challenge of maintaining the healthcare system as the proportion of older adults increases relative to the working-age population.

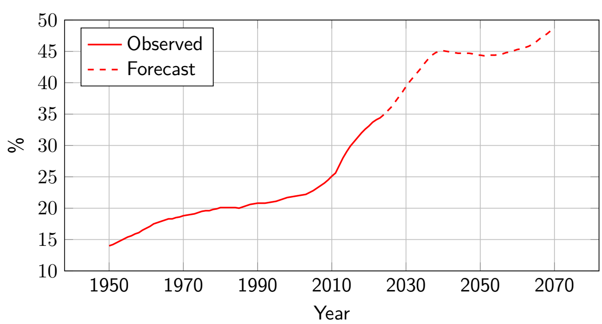


Figure 1.1: Grey pressure in the Netherlands (data source: SN).

Moreover, the country is experiencing a phenomenon known as 'double aging', characterized by both an increasing proportion of older adults and a rising average age. This phenomenon is marked by a growing proportion of individuals aged 80 and above, as shown in Figure 1.2. This age group typically exhibits the highest per capita healthcare utilization [47]. As a result, the increasing share of older adults and their healthcare needs impose a significant burden on the working-age population.

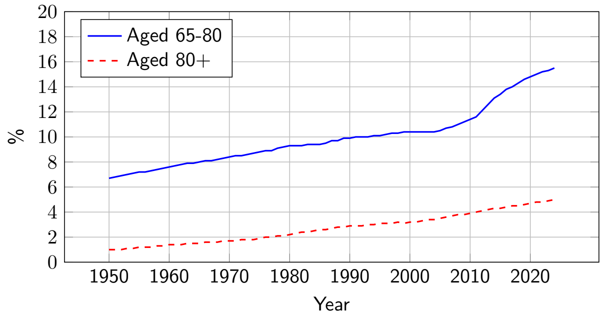


Figure 1.2: Double aging population (data source: SN).

The escalating grey pressure brings about several challenges. Firstly, there is a substantial increase in healthcare costs due to rising demand [113]. In the Netherlands, 45 percent of the healthcare budget in 2019 was allocated to individuals aged 65 and over, with nearly a quarter of this expenditure dedicated to dementia care [149]. Some countries have already opted to 'freeze' healthcare budgets to maintain affordability [37]. Additionally, there is a growing shortage of healthcare professionals, exacerbated by the increasing number of older individuals requiring care [55]. According to a study commissioned by the Dutch Ministry of Health, Welfare and Sports, the healthcare sector will need 135,000 workers by 2031, with a projected shortfall of 51,900 in nursing homes alone [133]. This shortage leads to capacity constraints across the healthcare system.

Furthermore, capacity limitations in one form of healthcare can have cascading effects on other forms due to difficulties in discharging patients when there is insufficient capacity in subsequent care settings. Therefore, capacity issues must be addressed from a system's perspective. In this thesis, we examine healthcare for older adults from a system's perspective. We analyze the *routes* taken by older adults through the healthcare system and focus on optimizing the decision-making process for older adults along their routes, here referred to as *routing*. Through our quantitative analysis and mathematical optimization, we aim to contribute to addressing the urgent and substantial challenges posed by the aging population.

1.1 DOLCE VITA Project

This thesis was developed within the framework of the NWO project named Data-Driven Optimization for a Vital Elderly Care System in the Netherlands (in short DOLCE VITA). This project represented a collaborative effort between mathematicians and geriatricians to model the healthcare system for older adults. The collaboration included Centrum Wiskunde & Informatica (CWI), Amsterdam UMC (the academic hospital of Amsterdam), Vrije Universiteit Amsterdam, and SIGRA [160], the umbrella organization for healthcare providers in Amsterdam. Furthermore, as end users, 23 healthcare organizations were actively involved, including the five major care organizations in Amsterdam, general practitioner organizations, and healthcare insurers. The consortium convened every six months to review progress and discuss developments.

The initial motivation for the project arose from a SIGRA report titled "Krakende Ketens" [159]. This report involved a detailed qualitative study of the healthcare system in Amsterdam, identifying critical bottlenecks within the system. Figure 1.3 illustrates the current state of the healthcare system in Amsterdam, highlighting various bottleneck transitions. The report [159] drew several key conclusions: (1) bottlenecks predominantly occur at points of care transition, (2) there is a misalignment between supply and demand, (3) insufficient coordination exists for integrated planning among care providers, and (4) the selection of care providers is primarily based on available capacity rather than patient preferences.

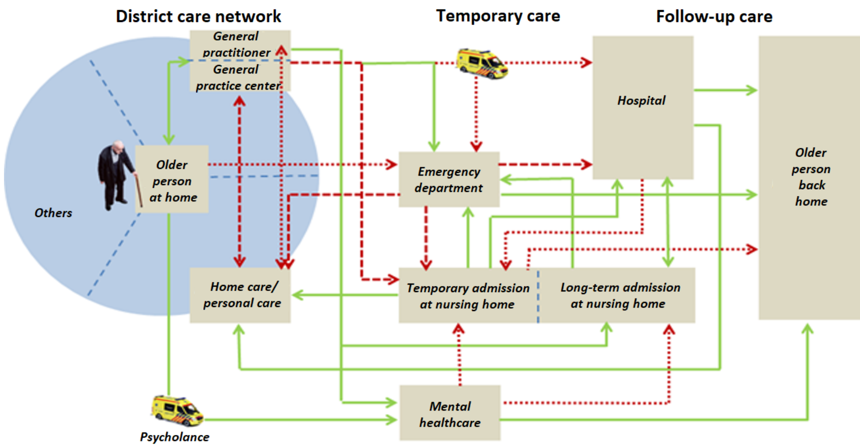


Figure 1.3: Overview healthcare system in Amsterdam, from [159]. Green line: no noteworthy bottleneck. Red dotted line: point of attention. Red solid line: large bottleneck.

The consequences of bottlenecks in the healthcare system were illustrated by a use case [159]; for a schematic representation, see Figure 1.4. This example involves an older woman with mild cognitive impairment who lives alone, receiving informal support from her children and regular home care. After falling down the stairs, her son takes her to the nearest

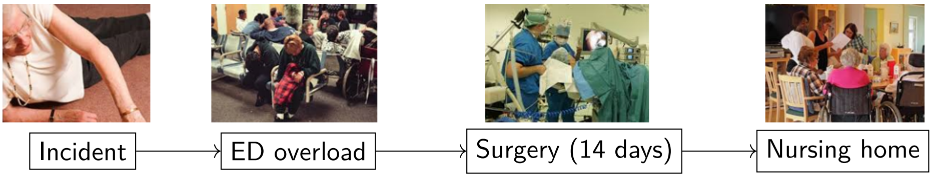


Figure 1.4: Example use case.

emergency department, which is overcrowded. Consequently, she has to wait six hours before a broken hip is finally diagnosed. She is then transferred to another hospital for surgery. Due to her cognitive impairment, she develops delirium, leading to a 14-day hospital stay with a significant decline in her functional abilities. As a result, she requires geriatric rehabilitation. However, due to her cognitive decline, she is unable to recover sufficiently and cannot return home, necessitating a wait for permanent admission to a nursing home.

This use case clearly illustrates how bottlenecks within the system can adversely affect patient outcomes, and how a malfunctioning care system can even generate additional demand. The report [159] only provided qualitative insights, lacking quantitative data. The objective of the DOLCE VITA project was to expand this analysis by quantitatively mapping the healthcare system using mathematical modeling and searching for system improvements through stochastic optimization.

1.2 The Healthcare System for Old Age

In the care system for older adults, we can distinguish between *short-term* care and *long-term* care forms. There are a few key differences in terms of the goal, duration, and age range of those who benefit from it. Short-Term Care (STC) is goal-oriented and needed after a condition in which the person is expected to make a full recovery back to their former relatively healthy and independent state. Long-Term Care (LTC) is necessary to help a person remain comfortable through the course of long-term illness, and is mostly useful for the elderly and those with progressive, degenerative diseases. LTC offers comprehensive treatment and a variety of amenities to maintain the quality of life for seniors.

The set-up of these care forms depends on the laws under which they are established. In the Netherlands, the care system for older adults is governed by three major laws, each addressing different aspects of healthcare and social services. These laws are the *Health Insurance Act (HIA)*, the *Long-Term Care Act (LTCA)*, and the *Social Support Act (SSA)*. We briefly elaborate on each of these laws.

The HIA is the cornerstone of the Dutch healthcare insurance system. Introduced in 2006, it mandates that all residents of the Netherlands must have basic health insurance, which covers essential medical care including visits to general practitioners, hospital care, and prescription medication. The HIA is designed to ensure universal access to healthcare services and is funded through a combination of premiums paid by individuals and contributions from employers [147].

The LTCA was implemented in 2015 to provide care for individuals who require permanent and intensive care due to chronic illness or disability. It covers long-term institutional care and severe cases of home care, ensuring that individuals who cannot manage daily activities independently receive appropriate support. The LTCA is primarily funded through contributions from residents and covers costs associated with nursing homes and similar facilities [145].

The SSA focuses on improving the quality of life for individuals with disabilities and the elderly by providing support services that help them live independently and participate in society. Introduced in 2007, the SSA covers services such as home assistance, adaptations to homes for accessibility, and support for social activities. Municipalities are responsible for the implementation and local administration of SSA services [146].

Although these laws coordinate all care forms for older adults in the Netherlands, these care forms also differ in facility type, as can be seen in Figure 1.5. The figure illustrates the three primary *locations* where care can be provided: at home, in the hospital, or in a nursing home.

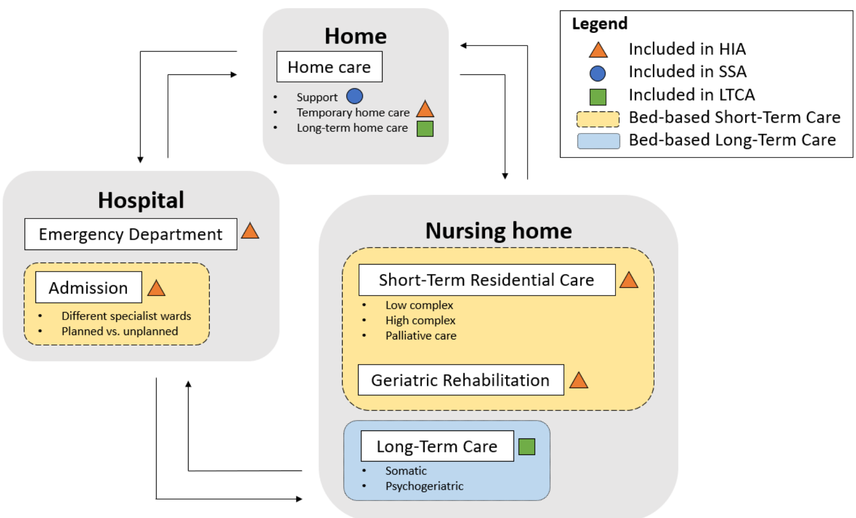


Figure 1.5: Schematic representation of the care system for old age in the Netherlands.

First, home care can be provided and financed under one of the three laws, depending on the care's intensity and objectives. Second, all hospital care is funded by the HIA. If a patient requires hospital admission to a specialist ward, this is classified as bed-based STC.

Third, nursing home facilities offer only bed-based care, which includes both STC and LTC, financed by the HIA and LTCA, respectively. It is important to note that the figure presented is a schematic representation; there are numerous exceptions across different regions of the Netherlands. Additionally, this analysis excludes care types that are beyond the scope of this research, such as care for older adults living in psychiatric institutions or care provided in prisons.

Figure 1.5 illustrates the fact that different locations provide distinct types of care, each regulated by different financial laws. As a result, modeling the entire care system is challenging due to the complex transitions between these various forms of care. In the following sections of this thesis, we provide a detailed analysis of Short-Term Care (part I) and Long-Term Care (part II).

1.3 Part I: Short-Term Care

For the Short-Term Care (STC) types, we focus exclusively on the primary categories of bed-based care: Geriatric Rehabilitation (GR), Short-Term Residential Care (STRC), and Hospitalization (HOS). Other forms of temporary care, such as LTCA crisis, which serves individuals with an LTCA designation who urgently require placement outside their homes, respite care, which offers short-term relief for primary caregivers, and bed-based care provided by the General Practitioner (GP), are not included in this analysis due to their relatively limited capacity. A brief overview of the GR, STRC, and HOS care types is provided below.

1.3.1 Geriatric Rehabilitation

Geriatric Rehabilitation (GR) care is a specialized field of geriatric medicine in which a geriatrician is responsible for the treatment. This care is multidisciplinary and includes services such as physical therapy, occupational therapy, and psychological care. It is primarily provided to frail older adults with the goal to eventually return home. Usually, GR care is needed after a hospitalization, e.g., after surgery for a fracture.

The GR care is financed by the HIA and is documented using a Diagnostics-Treatment Combination (DTC) healthcare trajectory. The DTC serves as a framework to determine the necessary type and amount of care for a treatment and its associated costs. For GR, the DTC trajectories provide an estimated Length-of-Stay (LoS) for each rehabilitation period, such as after surgery. The duration of a DTC is limited to a maximum of 120 days, whilst the average LoS of GR is 43 days [182]. In 2019, 53,000 persons received GR care, of which the majority had to recover after a fracture (30%), and second most after a stroke [182].

1.3.2 Short-Term Residential Care

Short-Term Residential Care (STRC) refers to a temporary stay in a nursing home for individuals who are unable to care for themselves at home at that time, with the goal of eventually returning home. Unlike medical specialist care provided during hospitalizations or GR, STRC involves short-term medical care aimed at helping vulnerable individuals recover and return to their homes. STRC can be classified into different levels of intensity: *high-complex* care, *low-complex* care, and *palliative* care. The distinction between high- and low-complex care lies in the complexity of the care required, while palliative care is provided at the end of life.

STRC is a relatively new type of care that was introduced in 2017. In its first year, 33,100 people received STRC, with the largest proportion being individuals aged 85 and older [179]. Of these, about 50% received high-complexity care, approximately 30% received low-complexity care, and around 20% received palliative care [179]. The number of people receiving STRC has increased over time, reaching 37,170 users in 2022 [199].

Note that the term STRC refers directly to the Dutch implementation of the more general Intermediate Care (IC) concept. IC is defined as care for individuals who require care more than room and board but below medical specialist care in hospitals, hence 'intermediate' [118]. Therefore, the terms STRC and IC are used in this thesis to describe this type of care services respectively specifically for the Netherlands or in general.

1.3.3 Hospital

Hospitalizations provide necessary care following an acute event, such as a severe injury or sudden illness. In addition to unplanned admissions due to emergencies, hospitalizations can also be planned for specific treatments or surgeries. After discharge, older adults often require additional care services, such as GR, to aid in their recovery and facilitate a return to their normal daily activities.

Hospital departments that commonly provide care to older adults include surgery, cardiology, and orthopedics, where patients receive specialized medical interventions tailored to their acute or chronic health conditions. For example, older adults may require surgical procedures due to fractures, cardiology interventions for heart conditions, or orthopedic care for joint replacements. On average, approximately 300,000 individuals aged 65 and older, accounting for about 9% of this population group, experience unplanned hospitalizations each year between 2017 and 2021 [178].

1.4 Part II: Long-Term Care

Long-Term Care (LTC) can be provided at home (home care) or at a nursing home location (nursing home care). The current policy is to stimulate persons to live longer at home, since this is considered less expensive. However, some recent research shows that if persons live longer at home, more costly hospital admissions will be needed [17]. To illustrate our LTC system, we will now further elaborate on the main long-term care forms.

1.4.1 Home Care

Home Care (HC) encompasses personal care and nursing services provided at home. This type of care may be required following hospitalization or due to chronic illness. Typically, HC is temporary, with the necessity for care being periodically reassessed. The primary goal is to enable individuals to live at home for as long as possible.

In 2017, approximately 550,000 individuals in the Netherlands received HC, with the majority of recipients being between the ages of 75 and 84 years [180]. Notably, within the 85+ age group, around 40% utilized HC, and approximately 56% of these individuals required somatic care for more than three months. On average, individuals aged 85 and older received 5.8 hours of care per week, with an average duration of six months [180]. According to the Dutch Healthcare Authority (DHA), there was an increase in the number of HC recipients from 2018 to 2022, with 570,000 to 590,000 individuals receiving care during this period [125].

1.4.2 Nursing Home Care

Nursing Home Care (NHC) is defined by the ongoing need for support and services throughout an individual's life due to aging-related conditions or disabilities. Within NHC, various care profiles are established based on the type and intensity of care required. This includes care specifically tailored for both older individuals facing aging-related issues and as well as those with disabilities; the focus here is on the former.

To access NHC services, an official assessment and indication from the Center for Indication of Care Needs (CICN) is mandatory. Once an indication is granted, the healthcare insurance office coordinates the allocation of care, arranging placements either in a nursing home or through home-based care services provided by various organizations.

The NHC profiles related to 'nursing and personal care' are classified under various "VV" profiles, each tailored to meet the specific needs of the recipients based on the intensity and type of care required. Generally, lower VV profiles correspond to less intensive care needs.

Firstly, we have VV 4, VV 6, and VV 8. These profiles are primarily for individuals requiring Somatic (SOM) care. VV 4 is designed for individuals with less severe somatic issues and can often be provided at home. In contrast, VV 8 is intended for individuals with severe somatic problems that necessitate admission to a nursing home. Secondly, we have VV 5 and VV 7. These profiles focus on Psychogeriatric (PG) care, typically for individuals with dementia. VV 5 is intended for general PG care, while VV 7 is specifically designed for those with dementia who also exhibit significant behavioral issues. Lastly, the VV 9b profile is specific to rehabilitation within a nursing home setting. It applies to individuals who either already have an existing NHC indication or will require one following their rehabilitation.

In 2017, approximately 223,000 older adults in the Netherlands received NHC, with about 92% of this care being provided under one of the VV profiles, primarily at nursing home facilities. Most nursing home residents required care for dementia, categorized under profile

VV 5 [180].

1.5 Overview of the Thesis

The thesis starts with Chapter 2, aiming to gain insight into the routes over the healthcare system for older adults. This study investigates the longitudinal patterns of healthcare utilization among older adults to enhance care coordination and optimize patient journeys. We analyze comprehensive healthcare usage data for all residents aged 65 and older in the Netherlands, covering the period from 2017 to 2019, obtained from microdata of Statistics Netherlands (SN) [39]. The study focuses on uncovering the most common patient journeys by categorizing patients based on age and medication use. We use process mining techniques to generate visual representations of the connections between different forms of care. By identifying these patterns, we provide valuable insights into improving healthcare delivery for older adults, particularly in terms of reducing unnecessary transitions and enhancing the coordination of care across various healthcare settings. This study is based on [59].

Thereafter, we focus on a part of the healthcare system in part I: Short-Term Care. Chapter 3 aims to analyze the characteristics of older adults admitted to STRC, a recent bed-based care concept in the Netherlands designed to prevent nursing home admission. We will examine patient profiles using data from a national STRC database, covering admissions from 2018 and 2019. In addition to patient characteristics, the study explores organizational differences between STRC facilities across the country. To achieve this, we conduct a cross-sectional email survey targeting all STRC facilities in the Netherlands. The survey collected information on various aspects of care delivery, such as the location of the wards, the number of beds, and the availability of medical services. By combining patient data with organizational insights, this study seeks to better understand the structure and outcomes of STRC, ultimately contributing to the development of care pathways that support older adults in need of reablement, observation, diagnostics, and palliative care. This study is based on [171].

In Chapter 4, we address the prolonged waiting times for IC in the Netherlands, which often lead to unnecessary and costly hospital admissions. We propose alternative policies to improve IC and estimate their potential effects on waiting times, hospitalization rates, and patient placement numbers. Using data from older adults who received IC in Amsterdam in 2019, we will analyze patient inflows, outflows, and characteristics. A process map of the main pathways into and out of IC is created, and a Discrete Event Simulation (DES) model is developed. This simulation allows us to evaluate various policy changes and their impact on the efficiency of the IC system. By focusing on real-life data from Amsterdam, our study seeks to provide actionable insights into optimizing IC processes, ultimately reducing waiting times and preventing unnecessary hospitalizations for older adults. This study is based on [12].

Then, we go on with Long-Term Care in part II. Chapter 5 addresses the challenge of excessive waiting times and patient abandonments in LTC systems, which are increasingly strained by the rapid aging of the population. These issues partly arise from inefficiencies in the allocation of patients to geographically dispersed care centers, each with its own preferences and levels of flexibility. To tackle this, we propose a novel and easy-to-implement method for optimally allocating patients to nursing homes. Our approach balances the trade-off between minimizing waiting times and accommodating individual patient preferences and flexibility. By focusing on optimizing patient placement, our method aims to enhance the efficiency of LTC systems, ensuring that patients are matched with suitable care facilities more quickly and in alignment with their needs. This approach is intended to support the development of patient-centered and sustainable LTC solutions. This chapter is based on [11].

Chapter 6 focuses on reducing long waiting times for nursing homes by implementing an advanced, preference-based allocation model for older adults. The research evaluates the performance of this model through simulations applied to three case studies in the Netherlands, encompassing urban regions like Amsterdam and Rotterdam, as well as a rural area in Twente. Data on nursing homes and their capacities were collected for each case study, and a set of preference profiles was developed to capture waiting time preferences and flexibility among older adults. The study also outlines guidelines for implementing the model in practice, taking into account the roles of all stakeholders involved. A simulation was conducted to compare the current practice with the proposed allocation model, focusing on outcomes related to waiting times and preferences. The goal is to demonstrate that this model can improve the efficiency of nursing home placements while better aligning with individual preferences. This study is based on [10] and [11].

Finally, in Chapter 7, we address the challenges in matching persons with severe mental illness to suitable residential facilities, a process complicated by both restrictions and client preferences. Traditional placement policies often prioritize easy-to-place clients, resulting in longer waiting times for those who are harder to place. To tackle this issue, we propose the Knapsack-Based Routing (KBR) model, designed to allocate clients to server pools more effectively, while considering both resource constraints and individual preferences. Our approach introduces a split-horizon method, which employs an integer quadratic programming technique to factor in both immediate and long-term impacts of scheduling decisions. The goal is to develop a robust scheduling policy that ensures a fairer distribution of waiting times across different client classes. This model is intended to enhance the placement process in residential facilities and can be applied to other service operations that involve complex routing and resource requirements, such as in healthcare and call centers. This chapter is based on [9].

2

Pathways over the Healthcare System

Based on:

de Boer, T. R., Arntzen, R. J., Bekker, R., Buurman, B. M., Willems, H. C., & van der Mei, R. D. (2025). Process mining on national healthcare data for the discovery of patient journeys of older adults. *Journal of the American Medical Directors Association*, *26*(1), 105333.

Abstract

Objective: Understanding the longitudinal patterns of healthcare utilization among older adults is crucial for designing effective patient journeys and enhancing care coordination across settings. This study aims to uncover the most common patient journeys of older adults.

Method: Data were sourced from Statistics Netherlands, encompassing all residents aged 65 or older as of January 1, 2017. Healthcare usage declarations from various care settings during 2017 to 2019 were included. Patient journeys were exclusively selected if their initiation points were certain. Data underwent rigorous preprocessing, merging, and filtering to create a single event log file suitable for Process Mining. Patients were categorized by age and medication use, and differences in patient journeys were analyzed. Process Mining techniques generated visualizations illustrating the connections between care forms and the impact of changes in one form on others.

Results: The study included 3,177,203 individuals aged 65 and older, with 44% experiencing one or more patient journeys totaling 2,469,663 journeys in the period of 2017-2019. Most care journeys for older adults were simple and short. The top ten most frequent journeys had four or fewer care forms, with 95% of journeys for the 65+ population and 90% for the 85+ population having four or fewer care transitions. Long-term care forms, such as home care, personalized care, and nursing home care, accounted for the majority of time spent in the system.

Conclusion: This pioneering study shows that most older adults tend to have a straightforward healthcare need, often involving the emergency department and hospitalizations. However, a smaller group among the population requires more complex and prolonged care, especially in the 85+ population. Reducing the number of transitions for this population, although impacting less people, might result in increased efficiency on the overall system.

2.1 Introduction

The demographic challenges posed by an aging population underscore the urgency of gaining a deeper understanding of healthcare usage among older adults. Such insights can be derived from Process Mining (PM), a method that employs advanced algorithms to analyze large volumes of data. PM analyses are typically represented as visualizations of patient journey patterns, illustrating the primary forms of healthcare utilized and the sequence in which they are accessed. This approach helps us to understand how patients navigate the healthcare system. For instance, PM can reveal a typical journey where a patient initially visits the emergency department, is subsequently admitted to the hospital, and finally receives home care.

To date, PM in healthcare has primarily been applied to smaller, more narrowly focused datasets, such as those pertaining to individual hospitals [110, 195]. However, PM has not yet been employed on a system-wide scale in healthcare. Achieving this broader application requires a comprehensive dataset. Statistics Netherlands (SN) collects data on healthcare claims from all residents of the Netherlands [39], enabling a holistic mapping of the healthcare system.

In this chapter, we aim to address the following two research questions:

1. What are the most frequently occurring patient journeys within the healthcare system among older adults?
2. Is there a difference in healthcare usage between various age groups and among groups with different levels of cumulative drug use?

The hypothesis is that older adults have long and intricate patient journeys, since most research often reports either long-term care pathways [87] or depicts the older adults as a frail population with recurrent healthcare usage [46]. It is expected that the length and complexity of these journeys will only increase as the age or medicine use of the subpopulations increases, since age [14] and medicine use [150] are both shown to correlate with an increase in healthcare usage and frailty.

2.2 Methods

This section covers the dataset and the methodology applied to analyze the data. Throughout, we use the following definition of a *patient journey*: A sequence of transitions between different forms of care, allowing for a maximum gap of four weeks without professional healthcare. *Healthcare usage* is defined as older adults utilizing professional healthcare, ranging from personal home care to hospital care. This four-week limit is based on research indicating that healthcare events within this timeframe are often connected [185].

The data contains healthcare usage from 2017 until 2019. Since it is possible that journeys have already started before 2017, the choice was made to focus only on journeys of which we are certain that at least four weeks of no care predate the earliest healthcare form, meaning only care journeys after February 1, 2017, are considered. It is certain that these journeys are not continuations from journeys that have started before 2017 and, therefore, we are certain to capture the beginning of the healthcare pathways. These new journeys are characterized by their number of care transitions (change of care form utilized by the patient), the used healthcare types and the time spent per journey.

2.2.1 Data

The dataset utilized in this study originates from multiple healthcare datasets provided by SN [39]. Permission was obtained to access the microdata via a secured portal. This study falls outside the scope of the Medical Research Subjects Act due to the complete anonymity of the dataset, the retrospective nature of the study, and the public availability of the data within SN.

The dataset includes the following forms of care: Home Care (HC), Personal Care (PC), Long-Term Care (LTC) at home or in an institution (respectively Nursing Home at Home (NHH) or Nursing Home (NH)), Geriatric Rehabilitation (GR), Short-Term Residential Care (STRC) (a distinction is made between non-palliative and palliative, respectively denoted by STRC and Palliative Short-Term Residential Care (STRC-pall)), Emergency Department (ED) visits and Hospitalization (HOS). No Care at Home (HOME) is used to define that a person used no declared care. As the different care forms are paid by different laws, rigorous preprocessing of the datasets was necessary to create a single event file suitable for PM [33].

The dataset comprises data from 3,177,203 individuals aged 65 or older on January 1, 2017. Data from 2017 to 2019 was selected to reflect the healthcare system, due to anomalies related to COVID-19 in 2020 [120] and the healthcare reform in 2015 [190]. Each row contains: an activity, a start date and an end date. Rows combined are seen as patient journeys; an example of such a patient journey could be the following: a patient has an ED, a HOS and returns home with PC. This journey then counts three care transitions and forms. The duration of this journey is then the length from start date of ED until the end date of PC. In addition, persons could have multiple forms of care at the same time. Periods with multiple care forms were handled by removing overlap based on a ranking in which the most intense care form remains. The ranking is based on intensity in the following order: Death, ED, HOS, GR, STRC, NH, NHH, PC and HC.

2.2.2 Process Mining

The aim of PM [175] is what is called process discovery [174], i.e. discovery of processes or paths in a system. This is done by analyzing log files using the R-library named bupaR [33]. We first generate a table containing the number of persons and patient journeys of each sub-population, after which we generate a table with the most frequently occurring patient journeys of each subpopulation. The journeys are then examined further, as we take a closer look at the distribution of the complexity of these journeys using a histogram and the time length using Lorenz curves. Finally, process maps are generated, which can provide a picture of the complexity of the taken care journeys of each group. These process maps are set to have a coverage of 0.8, meaning that 80% of the journeys are visualized. Increasing the coverage to 1.0 often creates maps containing a superfluous amount of flows, so-called ‘spaghetti maps’ [173], which are usually too detailed to be informative.

2.2.3 Age and polypharmacy

Finally, the population can be split based on age and drugs use. Specifically, the age ranges 65-74, 75-84 and 85+ are distinguished, where the age of the person on January 1, 2017, is used. Similarly, a distinction is made based on drugs use of a person in 2017, as drugs use can be seen as a proxy for frailty [150]. The population is split in persons that use 0-4 distinct types of medicine, 5-9, or 10+. These sub-groups enable comparisons between different demographic and medication profiles, with sample sizes of 100,000 persons per sub-population facilitating meaningful comparisons. Comparisons will be done between the 65+ and the 85+ population, results for other sub-groups are provided in the appendix.

2.3 Results

First, we give a general overview of the data, then we examine the care transitions and time distribution, and finally, we present the process maps.

2.3.1 Population data

Table 2.1 provides an overview of the population used in this analysis. The majority of the population is relatively young: 58% is aged between 65 and 74, and has little drugs use, only 21% has ten or more drugs. The fraction of persons with a care journey increases with age and drugs use, but quite surprisingly, the older sub-population of 85+ consists of fewer patients with a care journey (44% compared to 52% in the patient group 75-84 years of age). This can be explained by the fact that many older adults aged 85 or older were already in a care journey in January 2017. In this chapter, these journeys are omitted, which can be seen in the last row of Table 2.1. We also see that in terms of mean journey length the sub-population of 85+ differs most from the complete population, with an average of 2.45 care transitions compared to 2.06 overall.

Figures 2.1 and 2.2 illustrate the ten most frequent patient journeys for the 65+ and the 85+ populations, respectively. In Appendix 2.A.1, the ten most frequent patient journeys of all (sub-)populations are shown in Table 2.2. In both the figures and tables, many of these journeys involve less than five care transitions. The most frequently occurring journey for each (sub-)population is in all cases a journey containing only one healthcare form, often the journey consisting of an ED visit and afterwards returning HOME. This can be seen in

		Age			Medicin use		
	All	65-74	75-84	85+	0-4	5-9	10+
Number of persons (%)	3,177,203 (100%)	1,816,172 (58%)	969,814 (31%)	357,172 (11%)	1,381,582 (44%)	1,122,587 (35%)	673,034 (21%)
Number of patient journeys	2,469,663	1,241,416	936,401	273,261	605,638	948,280	915,745
Persons with a journey (%)	44%	39%	52%	44%	30%	49%	64%
Mean number of journeys per person	1.78	1.73	1.86	1.75	1.48	1.72	2.14
Mean journey length	2.06	1.89	2.18	2.45	1.81	1.96	2.33
Number of excluded patient journeys	612,470	150,243	236,985	223,212	160,857	189,503	262,110

Table 2.1: Population size (on January 1, 2017) and frequently occurring journeys.

Figure 2.1, which shows that this specific journey occurs 24,350 times in a sample of 100,000 persons. However, for the 85+ sub-population, the most frequently occurring journey is patients having PC and returning to HOME, also seen in Figure 2.2, which occurs 12,008 times in our sample of 100,000 persons. The importance of ED and the HOS is evident in the top-three journeys of the complete 65+ population, as they can be described by only the HOS and ED. This can also be seen in other top-ten journeys of other sub-populations.

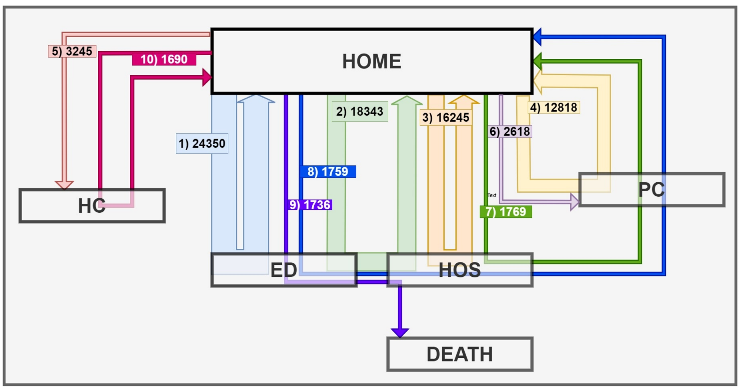


Figure 2.1: The ten most frequently occurring patient journeys for 65+ (taken from a sample of 100,000 persons). It shows how often the specific journeys occur within the sample.

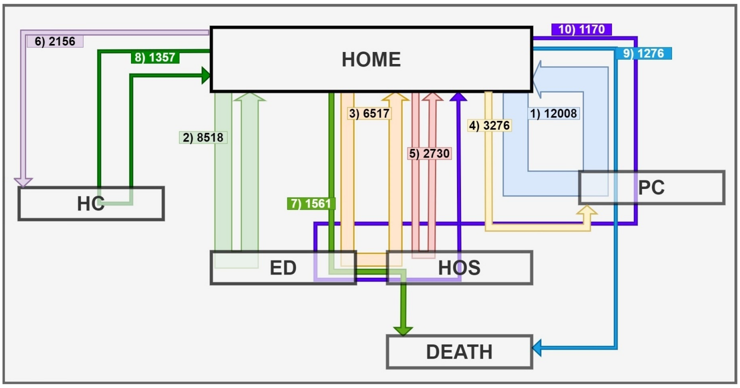


Figure 2.2: The ten most frequently occurring patient journeys for 85+ (taken from a sample of 100,000 persons). It shows how often the specific journeys occur within the sample.

2.3.2 Care transitions and time distribution

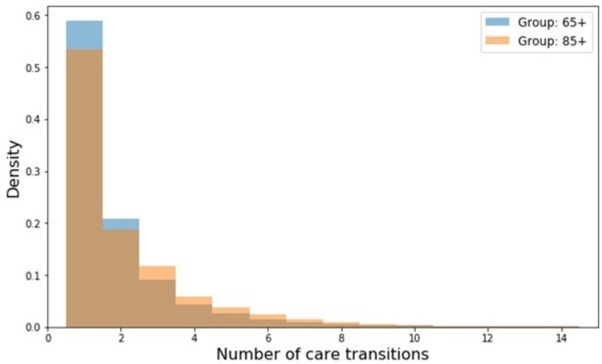


Figure 2.3: Histogram of the number of care transitions: a comparison between 65+ and 85+.

Now we zoom in on the care transitions and time durations. When looking at the distribution of the number of care transitions per journey, we observe that by far most care usage of older adults is simple and short, see Figure 2.3. Both the complete population and the older subpopulation exhibit a right-tailed distribution, with 59% of journeys for the 65+ population and 53% for the 85+ population involving only one care form. Additionally, 95% of journeys for the 65+ population and 90% for the 85+ population have four or fewer care transitions. Figure 2.4 visualizes the amount of time spent in the healthcare system using a Lorenz curve. In the 65+ group the 90% shortest journeys account for only 27% of the

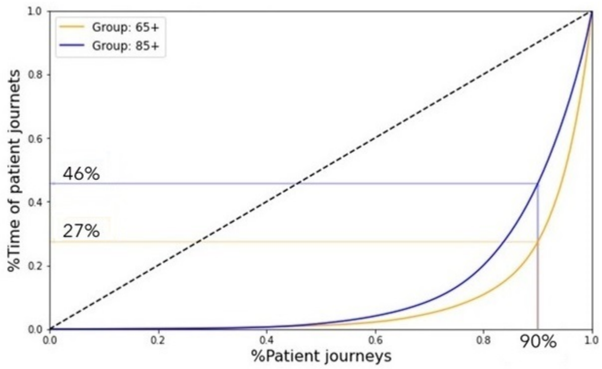


Figure 2.4: A Lorenz curve of patient journey time allocation between age groups 65+ and 85+. The curve illustrates the cumulative distribution of time spent across patient journeys, highlighting differences in healthcare utilization patterns among older adults.

total time. Hence, the 10% of the longest journeys account for 73% of the total time spent in the system, these 10% longest journeys consist mostly of LTC forms, such as HC, PC, LTC, nine of the ten most frequently occurring patient journeys of this 10% consist solely of HC, PC, LTC. The 85+ population shows that the time inequality is less severe.

2.3.3 Process maps

Figure 2.5 presents process maps for individuals aged 65 or older, while Figure 2.6 displays maps for those aged 85 or older. These maps offer a complete view of healthcare usage patterns within these age groups. They show that the majority of both groups have a short and straightforward healthcare journey (in line with Figures 2.1 and 2.2). However, by comparing 80% of population of 65+ and that of 85+, in respectively Figures 2.5 and 2.6, it can be seen that the older subpopulation has a more complex path. Notably, the process map for the 85+ subgroup, with the same coverage (0.8), exhibits more transitions and a less structured layout compared to the map for the 65+ population, indicating that healthcare usage among the older age group is less predictable and cannot be easily categorized into distinct pathways. This disparity is visually depicted by the "lasagna process" appearance of the map for the 65+ population, contrasting with the "spaghetti process" observed in the map for individuals aged 85 or older. Finally, in Appendix 2.A.2, the process maps are given for the other subpopulations.

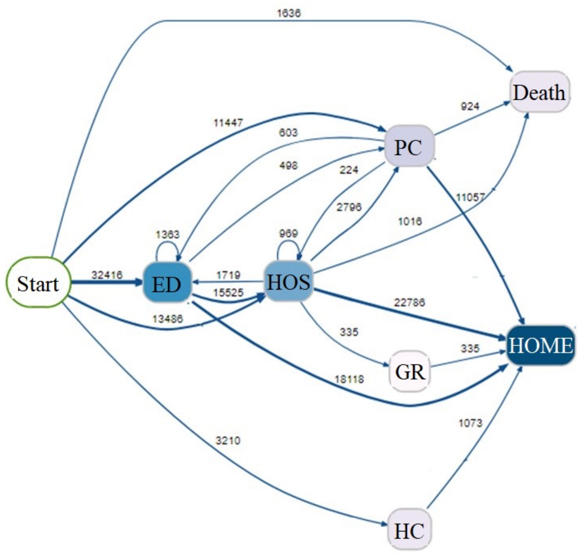


Figure 2.5: Process map of 80% of the most frequently occurring journeys of the 65+ population (taken from a sample of 100,000 persons), the number above the lines show how often that transition occurs in the sample, given an 80% coverage.

2.4 Discussion

This is the first study to identify patient journeys over a large-scale healthcare system, using Process Mining. It was found that most frequently occurring patient journeys of older adults contain either one or two and often not more than four care transitions. Similar to the simplicity it was found that most journeys span a relatively short amount of time, visible in our Lorenz curve. This confirms the previous observation that most older adults have simple and short healthcare journeys. Only a small part of the population covers more than half of the total time spent of all journeys, implying that this group requires more coordination in the system. Noteworthy is that this inequality diminishes for the 85+ population. This can be attributed to the fact that healthcare journeys that have started before 1 February 2017 are not included, and this subpopulation often utilizes long-term care, either at home or at an institution.

Another interesting observation is the consistent presence of ED visits or HOS in nearly all frequently occurring journeys across all (sub)populations. This underscores the important role of the ED and HOS as entry points into the healthcare system for older adults. Therefore, proper healthcare at the ED is crucial and strategies to avoid ED visits should be studied. Additionally, it was observed that HOS often followed ED visits, indicating the

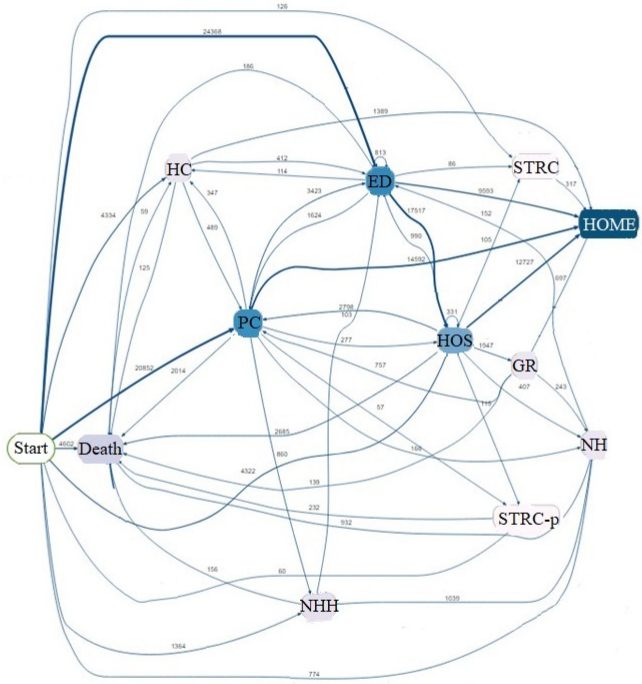


Figure 2.6: Process map of 80% of the most frequently occurring journeys of the 85+ population (taken from a sample of 100,000 persons), the number above the lines show how often that transition occurs in the sample, given an 80% coverage.

acute nature of HOS. Streamlining this transition is crucial given its prevalence in healthcare journeys, as inefficiencies in ED transfers can lead to inefficiencies in general [23].

Notably, the process maps of the older subpopulation 85+ depicted a more complex map than that of the population 65+, suggesting a greater diversity of healthcare journeys among older adults aged 85 and above. As the population of the Netherlands is aging more in the future [40], the strain on the system will rise further. More specifically, it is expected that the number of older adults aged 80 or above also increases over the coming years, for which we show that this group is harder to capture using a few transitions. This makes it more difficult to predict, anticipate and coordinate what journey an older adult aged 85 or older will embark on in the system. Simplifying these journeys could minimize efficiency lost in possible transfers, as transfers can be the source of a higher care demand, inefficiencies or dangers [154]. Integrated care for this population might be essential to limit the number of transfers [84]. For this sub-population, four out of the ten most frequently occurring journeys include PC, providing policy makers with a clear entryway into how to target this part of the population.

2.4.1 Strengths and limitations

This research distinguishes itself from previous research into patient journeys, by providing a scope on the complete healthcare system for old age, while previous papers often focus on either a small part of the healthcare system, e.g. the hospital, and only a small sample of persons [73, 80, 170]. Previous research on patient journeys of healthcare usage is often focused on designing the patient journey [20, 168], while this chapter focuses on first understanding what the patient journeys of older adults are, since the system should first be understood before adaptions can be made.

2.4.2 Future work

Some research opportunities fell beyond the scope of this chapter and could be topics of future research. Previous research done by the Dutch Healthcare Authority (DHA) [144] that focuses on the healthcare usage 100 days preceding death or nursing home admittance shows that these journeys are often different and visualize a different part of the population. This research approach is promising and could be examined further, especially given the importance of ED visits and HOS. Another opportunity for follow-up research is to include other groups of the population and a longer time window. Future research could also zoom in into different geographical regions, where the comparison between different regions with different types of healthcare is especially interesting, for example how does the proximity of ED [22] affect the patient journeys from patients.

2.5 Conclusion and Implications

In conclusion, this study employs PM on a dataset from SN to unravel the complexity of patient journeys of older adults in the Netherlands, addressing critical challenges in healthcare of older adults. The research aims to provide insights that can inform policy makers for more effective healthcare decisions, providing an overview into which and how different forms of care are used by the older adult's population of the Netherlands.

In summary, our study highlights how most older adults tend to have straightforward healthcare needs, often involving short visits to the ED or brief hospital stays. However, a smaller group among them requires more complex and prolonged care. This suggests that when designing or adapting new healthcare policies, focusing solely on the complex cases might overlook the majority who rely on simpler healthcare services. Therefore, it is important for policymakers to consider the broader picture, including routine emergency and hospital visits, to ensure that any changes made benefit the majority of older adults accessing healthcare. However, the 85+ population has a longer and more complex healthcare pathway. This implies that reducing the number of care transitions or journey length for only this group, might result in a larger effect on the system, due to the interconnectivity of various forms of healthcare that this group requires.

2.A Appendix

2.A.1 Most frequently occurring patient journeys

The ten most frequently occurring journeys of each sub-group can be found in Table 2.2.

	Activity				
Journey	**1st**	**2nd**	**3rd**	**4th**	**Number of journeys**
Group: all					
1st	ED	HOME			24350
2nd	ED	HOS	HOME		18343
3rd	HOS	HOME			16245
4th	PC	HOME			12818
5th	HC				3245
6th	PC				2618
7th	HOS	PC	HOME		1769
8th	ED	HOS	PC	HOME	1759
9th	ED	HOS	Death		1736
10th	HC	HOME			1690
Group: aged 65-74					
1st	ED	HOME			17862
2nd	HOS	HOME			12460
3rd	ED	HOS	HOME		9342
4th	PC	HOME			5202
5th	HC				1539
6th	ED	ED	HOME		1085
7th	HOS	PC	HOME		1040
8th	PC				869
9th	HC	HOME			830
10th	ED	HOS	PC	HOME	715
Group: aged 75-84					
1st	ED	HOME			16805
2nd	PC	HOME			13543
3rd	ED	HOS	HOME		10695
4th	HC				3361
5th	PC				2825
6th	HC	HOME			1512

Table continued from previous page

		Activity			
Journey	**1st**	**2nd**	**3rd**	**4th**	**Number of journeys**
7th	HOS	PC	HOME		1482
8th	ED	HOS	PC	HOME	1446
9th	ED	HOS	Death		1098
10th	PC	ED	HOS	HOME	939
		Group: aged 85+			
1st	PC	HOME			12008
2nd	ED	HOME			8518
3rd	ED	HOS	HOME		6517
4th	PC				3276
5th	HOS	HOME			2730
6th	HC				2156
7th	ED	HOS	Death		1561
8th	HC	HOME			1357
9th	PC	Death			1276
10th	PC	ED	HOS	HOME	1170
		Group: medicine use 0-4			
1st	ED	HOME			10870
2nd	HOS	HOME			6447
3rd	PC	HOME			5067
4th	ED	HOS	HOME		4780
5th	HC				1278
6th	PC				1048
7th	HOS	PC	HOME		577
8th	HC	HOME			567
9th	ED	HOS	PC	HOME	553
10th	ED	ED	HOME		480
		Group: medicine use 5-9			
1st	ED	HOME			18315
2nd	HOS	HOME			11792
3rd	PC	HOME			10059
4th	ED	HOS	HOME		9682
5th	HC				2608
6th	PC				2070

Table continued from previous page

		Activity			
Journey	1st	2nd	3rd	4th	Number of journeys
7th	HOS	PC	HOME		1284
8th	HC	HOME			1197
9th	ED	ED	HOME		1009
10th	ED	HOS	PC	HOME	994
Group: medicine use 10+					
1st	ED	HOME			20436
2nd	HOS	HOME			14387
3rd	ED	HOS	HOME		11255
4th	PC	HOME			6668
5th	HC				3452
6th	ED				2112
7th	HOS	PC	HOME		1894
8th	ED	HOS	PC	HOME	1645
9th	HC	HOME			1489
10th	PC	ED	HOME		1374

Table 2.2: The ten most frequently occurring patient journeys for the population aged 65-74 (taken from a sample of 100,000 persons).

In order to get more insight into the system's sub-groups, the distribution of care transitions and Lorenz curve for the time spent in the system distinguished in sub-groups are provided in Figures 2.7 and 2.8.

2.A.2 Process maps

The remaining process maps of other sub-groups can be found in Figure 2.9 to Figure 2.13 below.

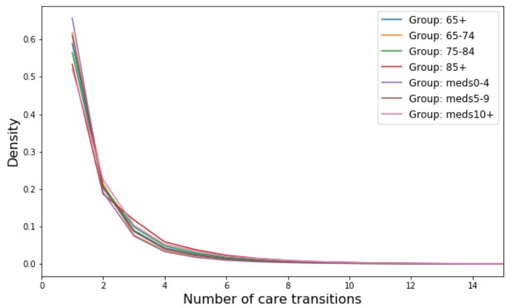


Figure 2.7: Exploring variations in care transition frequency distribution: a comparative analysis between (sub-)populations.

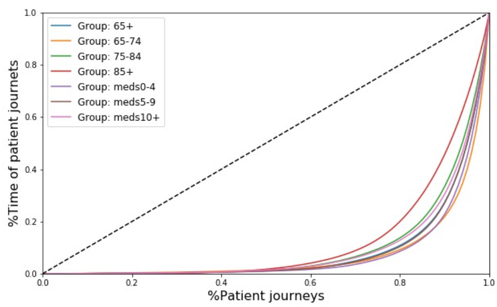


Figure 2.8: Interpreting the Lorenz curve: a comparative analysis of patient journey time allocation between (sub-)populations. The curve illustrates the cumulative distribution of time spent across patient journeys, highlighting differences in healthcare utilization patterns among older adults.

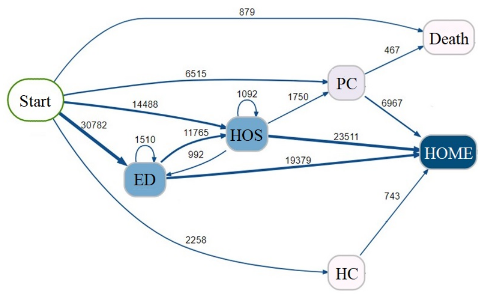


Figure 2.9: Process map of 80% of the most frequently occurring journeys of the 65-74 population (taken from a sample of 100,000 persons), the number above the lines show how often that transition occurs in the sample, given an 80% coverage.

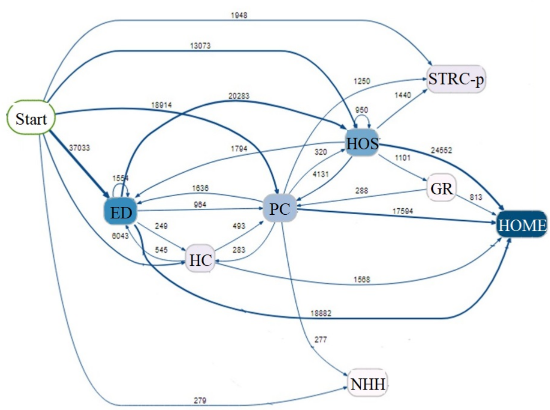


Figure 2.10: Process map of 80% of the most frequently occurring journeys of the 75-84 population (taken from a sample of 100,000 persons), the number above the lines show how often that transition occurs in the sample, given an 80% coverage.

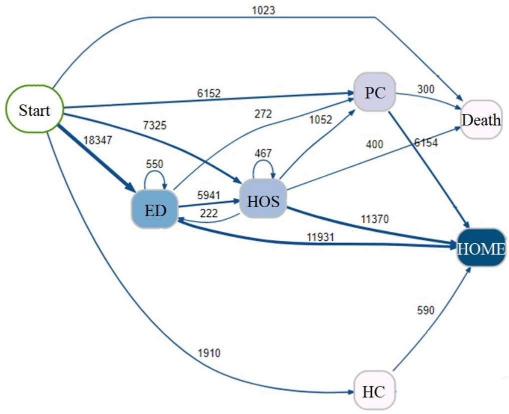


Figure 2.11: Process map of 80% of the most frequently occurring journeys of the 65+ population with medicine use between 0 and 4 (taken from a sample of 100,000 persons), the number above the lines show how often that transition occurs in the sample, given an 80% coverage.

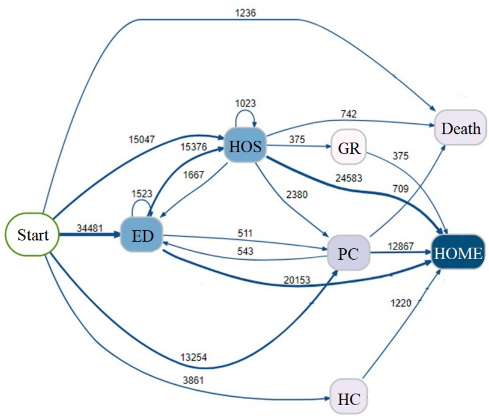


Figure 2.12: Process map of 80% of the most frequently occurring journeys of the 65+ population with medicine use between 5 and 9 (taken from a sample of 100,000 persons), the number above the lines show how often that transition occurs in the sample, given an 80% coverage.

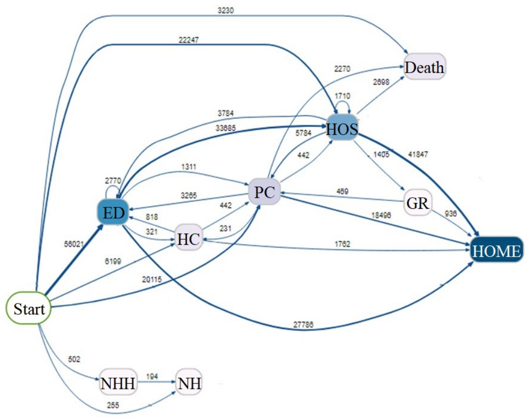


Figure 2.13: Process map of 80% of the most frequently occurring journeys of the 65+ population with medicine use higher than or equal to10 (taken from a sample of 100,000 persons), the number above the lines show how often that transition occurs in the sample, given an 80% coverage.

Part I

Short-Term Care

3

Survey on Organizational Features and Data Analysis of Intermediate Care

Based on:

van den Besselaar, J. H., Arntzen, R. J., MacNeil Vroomen, J. L., Hertogh, C. M. P. M., & Buurman, B. M. (2023). Short term residential care in the Netherlands: Patient and facility characteristics from a national database and survey. *Submitted*.

Abstract

Objective: Short-Term Residential Care (STRC) is a recent, bed-based care concept for older adults aimed at avoiding nursing home admission in the Netherlands. We aim to describe characteristics of patients admitted to STRC and the main organizational differences between facilities in the Netherlands.

Method: Patient characteristics were identified using a national STRC database (2018 and 2019). Organizational comparisons of STRC facilities were collected through a cross-sectional email survey sent to all facilities.

Results: Of the 68,682 older adults admitted to STRC, patients were mostly female (65%), living alone (64%), with at least 10 medications prescribed (60%). Of the 36,660 patients admitted in 2018, 43.4% died within 24 months and 25.5% are living in a nursing home. Of the 176 STRC facilities that responded to the email survey, 30.1% delivered care at an independent ward, 27.3% were within a geriatric rehabilitation care ward, and 33.5% at a ward in long-term care. The median number of beds was 8, with a range from 1 to 40. Most facilities admitted patients in evenings, nights, or weekends. Almost all wards employed registered nurses and paramedics.

Conclusion: Patients admitted to STRC have multiple medical problems, suggested by the high number of medications, and often have palliative care needs. Facilities providing STRC differ by the location of the ward, the number of beds, and frequency of medical rounds and multidisciplinary consultations. An important insight is the high mortality rate of older adults in the year after admission to STRC. It is important to develop care pathways to provide older adults with reablement, but also observation and diagnostics for long-term care and palliative care. Future research should focus on the outcomes of these care pathways and which older adults benefit from them.

3.1 Introduction

Over the last decade, European countries have reformed Long-Term Care (LTC) with a trend shifting from residential care towards home care and community care [164]. Residential care facilities have been reduced in number because of policies aiming for deinstitutionalization. To enable older adults to live independently, several countries implemented preventive and rehabilitation strategies [1, 88, 155, 164, 196].

In 2015, the Dutch government reorganized the health care system because of rising health care expenses [108]: LTC in nursing homes is from then on indicated only for patients who are in need of care 24 hours a day because of physical care needs and problems with self-management. The government wanted older adults to live as long as possible at home, with care close by. It was expected that older adults would need more often short term admissions for medical problems, with the goal to return home [85].

To address this care gap, a new care concept was implemented: Short-Term Residential Care (STRC). The target group consisted of older adults with general health problems that do not require admission to hospital for specialist care but that also cannot be treated at home. The goal was to enable older adults to return home and live independently in the community [165]. However, no national guidelines were provided for STRC and it soon appeared that the goal of STRC to support older adults in living longer independently at home was not met: in 2019, less than 50% of patients admitted to STRC returned home [181].

A previous qualitative study in three hospitals and eight nursing homes providing STRC in Amsterdam showed that older adults admitted to STRC had multiple complex problems in the medical, functional, psychological and social domains [172]. Advance care planning in the home situation is often lacking and patients often have a longer existing functional decline. This results in discharge to a nursing home instead of home, or even hospice care. Because of the complex problems, staff experienced that the tariff for STRC was inadequate to provide the right treatment and care. Further, that study found that the participating STRC facilities showed multiple organizational differences [172], especially in availability of qualified staff (nurses, paramedics and physicians), admission possibilities during the evening, night and weekend, and in the frequency of medical rounds and multidisciplinary consultations. As that study was only qualitative and focused on one city, it is unclear how STRC is provided on a national level and if other facilities have the same needs of a higher tariff, qualified personnel and advance care planning.

Therefore, the aim is to describe what the patient population of STRC looks like at a national level and whether the organizational differences between facilities also apply. If we know the patient population using STRC and the different organizational characteristics of facilities providing STRC at a national level, it is possible to compare outcomes such as discharge destination and length of stay between facilities in future research. This will enable the development of specific strategies to improve care. Using a national database of patients admitted to STRC and by distributing an exploratory national online survey, in this chapter, we address the following research questions:

- What are the characteristics of patients admitted to STRC in the Netherlands?
- What are the main organizational differences between individual facilities providing STRC in the Netherlands?

3.2 Methods

This was an observational, cross-sectional descriptive study. To describe the characteristics of patients admitted to STRC, we used the national database of Statistics Netherlands (SN). To describe the main organizational differences between individual facilities providing STRC, we distributed an online survey. Below, first we will describe the different tariffs of STRC, before elaborating on the study population, outcomes and statistical analysis of these two study methods.

3.2.1 Classification of STRC

STRC is bed-based care for general health problems with the goal to enable (older) adults to return home and live independently in the community. There are three different tariffs for STRC: regular, high-complex and hospice care. In STRC-regular the medical supervisor is the General Practitioner (GP) and only Activities of Daily Living (ADL)-care is provided by the facility. Paramedic treatment is according to the payment structure of the health insurer, as if the patient were at home. In the tariff for STRC-regular no treatment is included. STRC-regular is for patients who are not in need of specific paramedic treatment, but temporarily need more care than homecare can provide. For STRC-high complex the elderly care physician is the medical supervisor and up to 90 minutes per week of (para)medic treatment is funded. The elderly care physician is a medical practitioner who has specialized as a primary care expert in geriatric medicine [101] and has the goal to maintain or improve quality of life for older adults and chronically ill adults [167]. STRC-high complex is for patients who are not only in need of increased care, but also need (multidisciplinary) treatment or rehabilitation at a slower pace than in Geriatric Rehabilitation (GR). Third, STRC-hospice care is provided to patients in the last three months of their lives, the medical supervisor can be the nursing home physician or the GP. In STRC hospice care, 180 minutes of (para)medic treatment per week is funded.

3.2.2 National database

For the description of characteristics of the patients using STRC, a cohort was constructed through data linkage of multiple databases of SN [39]. We used a database of care expenses of all adults (18 years and older) in the Netherlands to select adults who used STRC in 2018 and 2019 to construct our study cohort.

The database of this study cohort was linked to databases with our main outcomes: age, gender, income, medication use, care utilization from 2017 to 2019, and cause of death.

The number of medications is registered as the total number of different drugs used in a year, excluding drugs prescribed during a hospital or nursing home admission. The database of care expenses contains the costs made in a whole year, but not the dates when these costs are made. This makes it possible to describe care utilization, but not the order in which care is used or the duration of the used care. Drugs are described using the ATC-code [130]. The cause of death is registered with the ICD-10 coding system [83].

3.2.3 National email survey

We developed a survey based on the different characteristics of STRC-facilities in Amsterdam [172]. We followed the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) checklist for the reporting of this observational research and the Checklist for Reporting of Survey Studies (CROSS) [157, 177]. The survey questions were composed with a physician elderly care with experience in STRC and GR and a nurse practitioner working in STRC. Further, the Dutch association of elderly care physicians (Verenso) [184], the National Health Care Institute (Zorginstituut) [200], the Dutch federation for long term care organizations (Actiz) [4] and the Dutch Healthcare Authority (DHA) provided feedback on the survey. The survey, consisting of 30 questions, was built in Limesurvey, an online survey program [107]. The questions were in Dutch and varied in closed and open form.

The survey addressed the following main areas through multiple questions: characteristics of the respondent; demographics of intermediate care facility; characteristics of the facility; specialization in specific patient groups; possibilities for patient admission; team organization and availability of paramedics; organization of care; use of guidelines in delivery of STRC (open question); discharge procedure; vision about target group and possible improvements for intermediate care (open question). The complete questionnaire is available upon request.

3.2.4 Participant recruitment for online survey

The survey was distributed between December 2019 and February 2020 amongst healthcare professionals working in STRC in the Netherlands. No national list of STRC providers exists. To reach healthcare professionals working in STRC, the survey was shared online on social media and newsletters of associations of health care organizations from nursing home physicians, nurses and general practitioners. Secondly, email addresses of providers of STRC were collected from www.zorgkaartnederland.nl, designed by the Dutch Patient Federation [137].

3.2.5 Sample size calculation survey

In total, 390 health care organizations in the Netherlands provide STRC [3], but it is unknown at how many locations this type of care is provided. For a representative sample of all organizations, we would need a sample size of 196 respondents working in different STRC-facilities for the survey to be representative [166].

3.2.6 Analysis

Data of characteristics of adults admitted to STRC are reported as mean with standard deviation (SD) and count with percentages as appropriate. Data was analyzed with R (Version R-4.0.5). For survey data, any doubles would be manually excluded by checking the timing of the respondents or if an incomplete response was present. Also surveys with an unknown facility or surveys that were filled in for multiple facilities at once were manually excluded. Results were presented in counts with percentages. Data from the survey was analyzed in SPSS (IBM SPSS Statistics for Windows, Version 26.0 Armond, NY: IBM Corp).

3.3 Results

First, we present the characteristics of older adults in STRC based on the national healthcare data, and thereafter the characteristics of the STRC organizations based on the national survey.

3.3.1 Characteristics of older adults in STRC

Table 3.1 shows the baseline characteristics of older adults admitted to STRC in 2018 and 2019. In total, 68,682 older adults were admitted to STRC in 2018 and 2019. Most patients admitted were female (64.7% in 2018, 65.0% in 2019), lived alone (64.1% in 2018, 64.2% in 2019) and about half of the patients had a low income (50.4% in 2018, 49.2% in 2019). Almost all patients use more than five different types of medication. In 2018 68.6% used ten or more different drugs and in 2019 60.8% used ten or more different drugs. Over a third of the patients admitted uses psychotropic drugs (2018: 38.6%, 2019: 36.5%).

Table 3.2 describes the longitudinal care utilization and mortality of older adults admitted to STRC in 2018. The table describes their care expenses in the year prior to admission (2017), the year of admission (2018) and the year after admission (2019). Almost all older adults visit the hospital in the year before and after the admission to STRC. 33.2% of patients using STRC are admitted to a long-term care facility in the year of STRC and 29.7% die in the same year. The year after admission, from the older adults that are still

* Low income: $\leq$140% of social minimum income, middle income: 140%-200% of social minimum income, high income: $\geq$200% of social minimum income.

** Based on ATC-code system, does not include medication provided in hospitals or under the Long-Term Care Act (LTCA) in nursing homes.

‡ Based on ATC Codes N05A (antipsychotics), N05B (anxiolytics), N05CD (benzodiazepine), N06A (antidepressants) and N06C (antidepressants in combination with psycholeptics).

	2018	2019
N	36,660	36,613
Male, N (%)	12,994 (35.3)	12,813 (35.0)
Age, mean (SD)	79.4 (10.8)	79.1 (10.9)
Income of household (N,%)*		
Low income	18,508 (50.4)	18,014 (49.2)
Middle income	11,048 (30.1)	11,098 (30.3)
High income	6,891 (18.7)	7,221 (19.7)
Living in institution	218 (0.6)	215 (0.6)
Unknown	76 (0.2)	65 (0.2)
Living situation, N (%)		
Living alone	23,502 (64.1)	23,505 (64.2)
Living together	12,864 (35.9)	12,828 (35.0)
Living in institution	218 (0.6)	215 (0.6)
Unknown	76 (0.2)	65 (0.2)
Number of drugs, N (%)**		
0	643 (1.7)	566 (1.4)
1-4	2,750 (7.5)	2,887 (7.9)
5-9	10,175 (27.8)	10,906 (29.8)
10-14	12,940 (35.3)	13,005 (35.5)
15-19	7,397 (20.1)	6,917 (18.9)
$>$20	2,755 (7.5)	2,332 (6.4)
Number of older adults using one or more psychotropic drugs, N (%)‡	14,119 (38.5)	13,361 (36.5)

Table 3.1: Adults admitted to STRC in 2018 and 2019 in the Netherlands.

alive, 25.5% are living in a LTC facility, 11.3% receive LTC at home and another 19.4% die. Of all patients admitted in 2018 to STRC, 15,887 (43.3%) died in the same or following year. The two major causes of death were cancer (38.8%) and diseases of the cardiovascular system (22.7%). In Appendix 3.A.1, Table 3.5 can be found describing all causes of death.

3.3.2 National survey characteristics of organization of STRC

We included 176 surveys for analysis. Figure 3.1 shows the flowchart of inclusion. Participants of the survey were mainly elderly care physicians (33.5%), managers (22.2%), nurses (15.9%) and staff working in front office (11.4%). The front office is responsible for the

* Outpatient clinic, emergency department and admission.

	2017	2018 (year of admission)	2019
	N=36,660	N=36,660	N=25,758
GP consultation, N (%)	34,242 (93.4)	34,305 (93.5)	19,818 (76.9)
Home care, N (%)	19,835 (54.1)	27,805 (75.8)	11,516 (44.7)
Hospital admission, N (%)*	34,564 (94.2)	35,331 (96.4)	22,631 (87.9)
Long term care at nursing home, N (%)	28 (0.8)	12,181 (33.2)	6,560 (25.5)
Long term care at home, N (%)	28 (0.8)	758 (2.1)	4,146 (11.3)
Death, N (%)	NA	10,902 (29.7)	4,985 (19.4)

Table 3.2: Longitudinal care utilization and mortality of older adults admitted to STRC in 2018 in the Netherlands.

planning of new admissions. 78.4% of the participants were directly involved in patient care at a STRC-ward. Full description on the profession of respondents is described in Appendix 3.A.2 in Table 3.6.

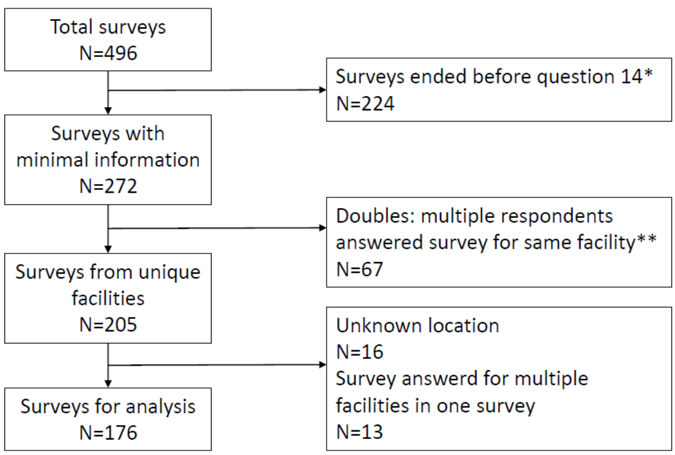


Figure 3.1: Flowchart of inclusion of surveys.

Table 3.3 describes the organizational structure and the contractual arrangement of the participating facilities. The participating facilities were distributed according to number of inhabitants over the 12 provinces of the Netherlands. 141 (80.1%) of participating wards provide STRC high complex, with a median number of beds of 8 (interquartile range 3-12, minimum 1, maximum 40). 85 facilities (48.3%) provide STRC regular care with a median number of beds of 3 (interquartile range 2-6, minimum 1, maximum 34). 83 (47.2%) facilities provide STRC hospice care, with a median number of beds of 3 (interquartile range 2-5, minimum 1, maximum 30). 153 (86.9%) of these facilities were Nursing Home (NH), Assisted Living Facilities (ALF) or both. In general, NH employs elderly care physicians who treat the residents, while in ALF the residents are treated by their GP. In general, ALF do not employ their own physicians and in practice multiple general practitioners treat residents. 92 (52.3%) of the facilities were smaller than 100 patients or residents. 79 (44.8%) also provided GR, 118 (67%) LTC for older adults with dementia and 131 (74%) LTC for older adults with a somatic disorder. The location of the STRC beds showed a large variation between organizations, but also within organizations: to the question where the beds were located organizations indicated multiple locations for one facility. Organizations have a combination of dedicated wards for STRC (30.1%), a shared ward with LTC (33.5%), GR (27.3%) or the beds of STRC are spread over the wards of the facility (19.3%).

In STRC high complex, the elderly care physician was in 87.2% the responsible physician, while in STRC regular the GP was in 90.6% the responsible physician. For STRC hospice care the elderly care physician is the responsible physician in 69.9% of the facilities, the GP in 22.9% and in 4.8% both were involved. Most STRC high complex (86.3%) and hospice care (89.2%) have the daily availability of a registered nurse. In STRC regular 59.0% of the facilities a nurse is available daily. 99.4% of STRC facilities have the availability of a physical therapist, 94.8% of an occupational therapist, 90.8% of a dietician, 82.1% of a speech therapist, 85.0% of a psychologist and 53.8% of a social worker (see Appendix 3.A.3, Table 3.7).

Table 3.4 shows the operating practices of the participating STRC facilities. On all items we saw a lot of variation: admission outside office hours, frequency of medical consultation and frequency of multidisciplinary team meetings. Specialization was uncommon for STRC facilities. Only 12 facilities described a specialization, these were high complex or hospice facilities. They specialized in patients with cognitive impairment, cognitive impairment and alcohol abuse, psychiatric problems, or tracheal cannula or peripherally inserted central catheters, or specialized in emergency admissions or as an observation unit.

Operating practice	**High Complex** N**=141**	**Low Regular** N**=85**	**Hospice Care** N**=83**
Admission of patients outside office hours, N (%)			
Yes, evening	98 (69.5)	44 (48.2)	43 (51.8)
Yes, night	63 (44.7)	23 (27.1)	25 (30.1)
Yes, weekend	97 (68.8)	46 (54.1)	45 (54.2)

* We included surveys which were completed until the 14th question to make sure we had enough information of the facility. Questionnaires which were ended before the 14th question were excluded.
** If multiple respondents answered the questionnaire for the same facility we included the questionnaire which was answered most of the questions. If there were multiple questionnaires completed, we included the first completed questionnaire.

Table continued from previous page

Operating practice	High Complex N=141	Low Regular N=85	Hospice Care N=83
No admission outside office hours	28 (19.9)	26 (30.6)	26 (31.3)
Don't know	9 (6.4)	10 (11.8)	8 (9.6)
Frequency of physician's visit, N (%)			
Weekly	80 (58.0)	11 (13.4)	56 (69.1)
Every two weeks	15 (10.9)	8 (9.8)	2 (2.5)
Every three weeks	0	1 (1.2)	0
Every month	1 (0.7)	1 (1.2)	0
When needed, regular moment	1 (0.7)	0	1 (1.2)
When needed, no regular moment	32 (23.2)	52 (63.4)	16 (19.8)
None	0	0	0
Other	3 (2.2)	9 (11.0)	6 (7.4)
Don't know	6 (4.3)	0	0
Frequency patient chart review, N (%)			
Weekly	98 (71.0)	14 (17.1)	49 (60.5)
Every two weeks	6 (4.3)	4 (4.9)	1 (1.2)
Every month	1 (0.7)	4 (4.9)	0
When needed, regular moment	2 (1.4)	2 (2.4)	4 (4.9)
When needed, no regular moment	16 (11.6)	38 (46.3)	12 (14.8)
None	1 (0.7)	2 (2.4)	1 (1.2)
Don't know	14 (10.1)	18 (22.0)	14 (17.3)
Frequency of multidisciplinary team meetings, N (%)			
Once per week	50 (36.0)	14 (16.9)	33 (40.7)
Once every 2 weeks	36 (25.9)	12 (14.5)	17 (21.0)
Once every 3 weeks	13 (9.4)	2 (2.4)	2 (2.5)
Once every month	9 (6.5)	4 (4.8)	5 (6.2)
Less than every month	7 (5.0)	7 (8.4)	7 (8.6)
When needed	8 (5.8)	1 (1.2)	1 (1.2)
None	0	32 (38.6)	0
Don't know	5 (3.6)	11 (13.3)	5 (6.2)
Other	4 (2.8)	0	0
Goal setting with planned discharge date?, N (%)			
Yes	120 (85.1)	60 (70.5)	NA
No	10 (8.2)	13 (15.3)	NA

Table continued from previous page

Operating practice	High Complex N=**141**	Low Regular N=**85**	Hospice Care N=**83**
Don't know	4 (2.8)	5 (5.9)	NA
Other	3 (2.1)	3 (3.5)	NA
Missing	4 (2.8)	4 (4.7)	NA

Table 3.4: Operating practices of facilities providing short-term residential care in the Netherlands participating in survey.

STRC facilities used guidelines for the comprehensive geriatric assessment, multidisciplinary consultations, GR, oncologic care and palliative care. Staff is missing guidance in the difference between GR and STRC: both admit frail older adults, but the tariff for GR is higher than for STRC. Some staff pointed out that STRC and GR should be combined or show a lot of similarities. Almost all staff pointed out that the tariff is not sufficient for reablement or observation and diagnostics in STRC. Also, a lot of the older adults are in need of LTC. Staff wanted guidelines for how to triage, for reablement, observation and for responsibilities and quality in STRC. Some staff pointed out they wanted their STRC beds to be clustered on a dedicated ward, instead of scattered over the facility.

3.4 Discussion

We aimed to describe the patient population using STRC in the Netherlands and describe the organizational differences using a national database and an exploratory survey.

STRC is mainly used by females and older adults who live alone. Over 60% of patients used more than ten different drugs in the year of admission, suggesting multimorbidity, and a third used psychotropic drugs. Further, patients are very vulnerable, since 43% of the patients died in the same year of admission or in the year after the admission to STRC. STRC is mainly delivered in NH and ALF. Facilities locate their STRC beds heterogeneously: there is almost an equal division of facilities that locate their STRC beds on a dedicated STRC ward, a ward with GR or a ward with LTC. A fifth of the participating organizations also indicated the beds for STRC are scattered over the different wards of the facility. Staffing is similar over the participating facilities with good availability of a registered nurse and paramedic care. The participating organizations differed in how often they provide medical and multidisciplinary consultations.

This study confirmed the results of our previous qualitative study of patient cases in Amsterdam. These also concerned mainly older adults living alone, with a lot of medical problems and often psychiatric diseases. Patients admitted to STRC have a comparable age and gender to patients admitted to community hospitals in the UK [64] and patients admitted to post-acute care in skilled nursing facilities in the US [189]. Of the patients in post-acute

N = **176**	
Facility, N (%)	
Nursing Home (NH)	100 (56.8)
Assisted Living Facility (ALF)	46 (26.1)
NH and ALF	7 (4.0)
Rehabilitation center	7 (4.0)
Hospital	10 (5.7)
Hospice	2 (1.1)
Care hotel	4 (2.3)
Number of residents or patients	
<51	37 (21)
51-100	55 (31.3)
>100	82 (46.6)
Don't know	2 (1.1)
Other types of care delivered at facility (multiple response)	
Geriatric rehabilitation care	79 (44.9)
LTC | psychogeriatric	118 (67.0)
LTC | psychiatric gerontology	34 (19.3)
LTC | somatic care	131 (74.4)
Respite care	35 (19.9)
Sheltered home	36 (20.5)
Home care	50 (28.4)
Province	
Noord-Holland	24 (13.6)
Zuid-Holland	37 (21.0)
Utrecht	9 (5.1)
Groningen	7 (4.0)
Friesland	10 (5.7)
Flevoland	5 (2.8)
Gelderland	23 (13.1)
Overijssel	15 (8.5)
Drenthe	5 (2.8)
Zeeland	6 (3.4)
Noord-Brabant	24 (13.6)
Limburg	11 (6.3)

Table 3.3: Organizational structure and contractual arrangement of facilities providing short-term residential care in the Netherlands participating in survey.

care in the US two third have eleven or more chronic conditions, which might reflect the high medication use of the Dutch population. Death rates in STRC are comparable to the one-year-mortality rate in community hospitals, however the number of patients dying of cancer in STRC is higher than of the community hospitals in the UK [64]. Death rates in intermediate care or community rehabilitation have been rising in the UK the last years. It has been advocated to integrate palliative care services for older people, as the complexity of cases being cared for in the community is increasing [8].

The organization of STRC is heterogeneous, which makes it difficult to compare outcomes between facilities. This is also the case for intermediate care facilities in the UK [112, 122]. Two studies that tried to evaluate the effectiveness of different programs of intermediate care in the UK failed, because they did not reach a sufficient sample size [21, 197]. Future research should focus on different care pathways of STRC and their outcomes. Staff indicated a higher tariff is needed and the development of care pathways for STRC. Future guidelines should focus on two patient categories: (1) *reablement* and (2) *observation for cognitive problems*. Reablement could be, as suggested by staff, located near geriatric rehabilitation, while older adults admitted for cognitive observation could profit from being located near LTC. The higher tariff in the reablement care pathway could be used for more physical therapy, while in the observation care pathway this could be used for psychological observation and support of social work. The description of these pathways with quality indicators will make it easier to compare outcomes between facilities. Further, palliative care should have an important role in these care pathways, since the high mortality rate of older adults admitted to STRC.

3.4.1 Strengths and limitations

Our study is the first national overview of the patients admitted and organization of care in short term residential care in the Netherlands. We objectified the differences and similarities of facilities providing STRC in the Netherlands, further we gave insight in the needs of staff in STRC. This study makes it possible to compare STRC to other types of intermediate care in the world.

A limitation is that the dataset of care costs does not provide admission dates: we were not able to calculate the exact one-year mortality, nor were we able to determine the order in which care was used during the year. Because the costs did not describe if an older adult was using STRC regular, high complex or hospice care we could not analyze the mortality rate of patients of the separate types of STRC. However, we expect the total amount of STRC hospice care to be small according to our survey results. Most of the facilities provide STRC high complex and these wards are also larger than wards of STRC low complex and hospice care. The high mortality rate is in this way not only explained by the hospice care admissions. Also, it was not possible to determine different patient characteristics between STRC regular, high and hospice care.

Because it is unclear how many facilities provide STRC in the Netherlands we do not know if our sample size is representative. We made an estimation of approximately 400 different locations in the Netherlands, which would mean we needed a sample size of 196 locations. However, the distribution of the different locations is representative for the number of inhabitants of the different provinces of the Netherlands. Also, we had a lot of incomplete surveys of general practitioners who also treat patients in STRC, which could reflect a selection bias of facilities.

This first national study on STRC confirmed a complex patient population and the large differences in operating practices such as medical and multidisciplinary consultations. An important insight is the high mortality rate of older adults in the year after admission to STRC. It is important to develop care pathways and guidelines that make it possible to provide older adults with reablement, but also provide observation and diagnostics for LTC and palliative care. Future research should focus on the outcomes of these care pathways and which older adults benefit from them.

3.A Appendix

3.A.1 Causes of death

The causes of death in the STRC facilities were analyzed using the SN national database. The results of this analysis are provided in Table 3.5.

Causes of death	**2018, N=10,902**	**2019, N=4,985**
Cause of death top 10 in 2018, N (%)		
Lung cancer	1,165 (10.6)	218 (4.5)
Heart failure	624 (5.7)	305 (6.1)
Breast cancer	322 (2.9)	100 (2.0)
Prostate cancer	310 (2.8)	103 (2.1)
Family history of stroke	299 (2.7)	138 (2.8)
Colon cancer	293 (2.7)	86 (1.7)
Dementia	287 (2.6)	355 (7.1)
Pneumonia	260 (2.4)	167 (3.4)
COPD	237 (2.2)	145 (2.9)
Fall	188 (1.7)	116 (2.3)
Top 5 cause of death 2018, major groups ICD		
C - Neoplasms	4,940 (45.3)	NA
I - Disease of circulatory system	2,335 (21.4)	1,239 (24.9)
J - Disease of respiratory system	845 (7.8)	1,279 (25.7)
F - Mental, behavioural and neurodevelopmental disorders	366 (3.4)	446 (8.9)

Table continued from previous page

Causes of death	2018, N=10,902	2019, N=4,985
K - Diseases of the digestive system	321 (2.9)	276 (5.5)
Top 5 cause of death 2019, major groups ICD		
C - Neoplasms	NA	1,239 (24.9)
I - Diseases of circulatory system	NA	1,279 (25.7)
J - Diseases of the respiratory system	NA	525 (10.5)
F - Mental, behavioural and neurodevelopmental disorders	NA	446 (8.9)
G - Diseases of the nervous system	NA	276 (5.5)
Cause of death, major groups ICD, N (%)		
A - Certain infectious and parasitic diseases	208 (1.35)	NA
C - Neoplasms	6,179 (40.2)	1,239 (24.9)
D - Diseases of the blood and blood-forming organs	338 (2.2)	1,279 (25.7)
E - Endocrine, nutritional and metabolic diseases	305 (2.0)	525 (10.5)
F - Mental, behavioural and neurodevelopmental disorders	812 (5.3)	446 (8.9)
G - Diseases of the nervous system	565 (3.7)	276 (5.5)
I - Diseases of the circulatory system	3,614 (23.5)	1,370 (8.9)
J - Diseases of the respiratory system	1,370 (8.9)	477 (3.1)
K - Diseases of the digestive system	477 (3.1)	163 (1.1)
M - Diseases of the musculoskeletal and connective tissue	487 (3.2)	441 (2.9)
N - Diseases of the genitourinary system	389 (2.5)	NA
R - Symptoms, signs, and abnormal findings	441 (2.9)	NA
W - External causes of morbidity	NA	NA

Table 3.5: Causes of death.

3.A.2 Profession of respondents

In the survey study, the profession of the respondents were obtained. The results are provided in Table 3.6.

3.A.3 Staffing and skill mix

In the survey study, the staffing and skill mix of the STRC facilities were obtained. The results are provided in Table 3.7.

* Plans the admission of older adults in STRC

	N=176
Profession N (%)	
Elderly care physician	59 (33.5)
General practitioner	4 (2.3)
Resident elderly care	2 (1.1)
Physician assistant	5 (2.8)
Nurse	28 (15.9)
Health care assistant	4 (2.3)
Physical or occupational therapist	2 (1.1)
Social worker	1 (0.6)
Manager	39 (22.2)
Front-office*	20 (11.4)
Account manager	5 (2.8)
Board member health care organization	3 (1.7)
Team coach	3 (1.7)
Advisor quality and care	1 (0.6)
Involved in direct patient care at ward	138 (78.4)

Table 3.6: Professions of respondents.

Profession/Question	**High Complex** N=141	**Regular** N=85	**Hospice Care** N=83
Responsible/supervising physician, N (%)			
Elderly care physician	123 (87.2)	8 (9.4)	58 (69.9)
General practitioner	15 (10.6)	77 (90.6)	19 (22.9)
Specialist elderly care or GP	1 (0.7)		4 (4.8)
Specialist elderly care or physician assistant	2 (1.4)		1 (1.2)
Don't know			1 (1.2)
Other			1 (1.2)
In charge of daily care (multiple response), N (%)			
Elderly care physician	96 (68.1)	5 (5.9)	50 (60.2)
General practitioner	13 (9.2)	41 (48.2)	17 (20.5)
Resident	20 (14.2)		12 (14.5)
Intern	21 (14.9)	1 (1.2)	6 (7.2)
Physician assistant	41 (29.1)	8 (9.4)	17 (20.5)
Nurse	60 (42.6)	48 (56.5)	40 (48.2)
Health care assistant	55 (39.0)	56 (65.9)	34 (41.0)

Table continued from previous page

Profession/Question	High Complex N=**141**	Regular N=**85**	Hospice Care N=**83**
Don't know	2 (1.4)	2 (2.4)	1 (1.2)
Other	1 (0.7)	2 (2.4)	1 (1.2)
Availability physician assistant, N (%)			
Daily	14 (9.9)		7 (8.4)
Weekly	40 (28.4)	7 (8.2)	20 (24.1)
Monthly	5 (3.5)	2 (2.4)	1 (1.2)
Consult	21 (14.9)	16 (18.8)	16 (19.3)
Never	54 (38.3)	51 (60.0)	32 (38.6)
Don't know	5 (3.5)	7 (8.2)	5 (6.0)
Missing	2 (1.4)	2 (2.4)	2 (2.4)
Availability skilled nurse level 5*, N (%)			
Daily	38 (27.0)	9 (10.6)	28 (33.7)
Weekly	31 (22.0)	21 (23.5)	16 (19.3)
Monthly	2 (1.4)	3 (3.5)	2 (2.4)
Consult	23 (16.3)	10 (11.8)	13 (15.7)
Never	37 (26.3)	33 (38.8)	20 (24.1)
Don't know	8 (5.7)	8 (9.4)	2 (2.4)
Missing	2 (1.4)	2 (2.4)	2 (2.4)
Availability skilled nurse level 4*, N (%)			
Daily	119 (84.4)	49 (57.6)	71 (85.5)
Weekly	15 (10.6)	16 (18.8)	7 (8.4)
Monthly	0	3 (3.5)	1 (1.2)
Consult	1 (0.7)	3 (3.5)	0
Never	1 (0.7)	3 (3.5)	0
Don't know	3 (2.1)	9 (10.6)	2 (2.4)
Missing	2 (1.4)	2 (2.4)	2 (2.4)
Availability community nurse, N (%)			
Daily	NA	4 (4.7)	NA
Weekly	NA	5 (5.9)	NA
Monthly	NA	3 (3.5)	NA
Consult	NA	16 (18.8)	NA
Never	NA	45 (52.9)	NA
Don't know	NA	10 (11.8)	NA

Table continued from previous page

Profession/Question	High Complex N=141	Regular N=85	Hospice Care N=83
Missing	NA	2 (2.4)	NA
Who is mainly providing ADL-care?, N (%)			
All levels	50 (35.5)	19 (22.4)	30 (36.1)
Skilled nurse level 5*	2 (1.4)	0	0
Skilled nurse level 4*	20 (14.2)	5 (5.9)	22 (26.5)
Registered level 3*	61 (43.4)	50 (58.8)	26 (31.3)
Health assistant level 2*	0	1 (1.2)	0
Other	2 (1.4)	2 (2.4)	2 (2.4)
Don't know	4 (2.8)	6 (7.1)	1 (1.2)
Missing	2 (1.4)	2 (2.4)	2 (2.4)
Availability of paramedics, N (%)	All levels together, total N=173		
Physical therapist	172 (99.4)		
Occupational therapist	164 (94.8)		
Dietician	157 (90.8)		
Speech therapist	142 (82.1)		
Psychologist	147 (85.0)		
Social worker	93 (53.8)		
Pharmacist	92 (53.2)		
Activity support	89 (51.4)		

Table 3.7: Staffing and skill mix in facilities providing short term residential care in the Netherlands.

* In the Netherlands, skilled nurses and registered nurse assistants can be trained according to different levels with different responsibilities.

4

Modelling and Simulation of Waiting Times in Intermediate Care

Based on:

Arntzen, R. J., van den Besselaar, J. H., Bekker, R., Buurman, B. M., & van der Mei, R. D. (2023). Avoiding hospital admissions and delayed transfers of care by improved access to intermediate care: A simulation study. *Journal of the American Medical Directors Association*, *24*(7), 945–950.

Abstract

Objective: The current waiting times for Intermediate Care in the Netherlands prohibit timely access, leading to unwanted and costly hospital admissions. We propose alternative policies for improvement of Intermediate Care and estimate the effects on the waiting times, hospitalization, and the number of patient replacements.

Method: For our case study, data was used of older adults who received Intermediate Care in Amsterdam, The Netherlands, in 2019. For this target group, in- and outflows and patient characteristics were identified. A process map of the main pathways into and out of the Intermediate Care was obtained and a Discrete Event Simulation was built. We demonstrate the use of our Discrete Event Simulation for Intermediate Care by evaluating possible policy changes for the real-life case study Amsterdam.

Results: By means of a sensitivity analysis with the Discrete Event Simulation, we show that in Amsterdam the waiting times are not a result of a lack in bed capacity, but are due to an inefficient triage and application process. Older adults have to wait a median of 1.8 days for admission, leading to hospitalization. If the application process becomes more efficient and evening and weekend admissions are allowed, we find that unwanted hospitalization can be reduced substantially.

Conclusion: In this chapter, a simulation model is developed for Intermediate Care that can serve as a basis for policy decisions. Our case study shows that the waiting times for healthcare facilities are not always solved by increasing bed capacity. This underlines the necessity for a data-driven approach to identify logistic bottlenecks and finding the best ways to solve them.

4.1 Introduction

Older adults admitted to the hospital are at risk of adverse health outcomes, such as infections, delirium, falls and even death [30, 35, 49]. Moreover, prolonged hospital stays caused by delayed transfers of care often lead to adverse events [106, 140] and high costs [82]. In response to this, during the last decade many governments implemented Intermediate Care (IC) with the aim to avoid hospital admissions and to support older adults in their recovery after hospital admission. Currently, inadequate access to IC hinders the potential benefits of the facilities to the care system [163]. If we know what IC capacity is needed and which policies perform best, we can prevent hospital admissions and the consequent high costs.

It is, however, not trivial to determine which policies are the most effective for the improvement of IC. A solid method that can be used to predict outcomes of policy changes is Discrete Event Simulation (DES)). With this approach, process characteristics are identified and modeled as a sequence of events in time. DES is applied to many healthcare settings, e.g., emergency departments [48] or intensive care [79]. However, little research has been conducted on simulation of IC facilities. To the best of our knowledge, no study models IC as a mid-chain care form in which avoidable hospital admissions are considered as well, which is a necessity for determining the system-wide policy effects.

In this study, we develop a DES model that can be used to evaluate which policy change is likely to be the most effective in the prevention of bed-blocking in hospitals and intermediate care. Using the model, we run a case study for the city of Amsterdam, The Netherlands. In this study, we address the following research question:

- How can the organization of IC be adjusted such that the number of hospital admissions, delayed transfers of care and patient transitions are reduced?

4.2 Method

4.2.1 Background

IC is defined by an international panel as a broad range of time-limited services, from crisis response to support for several weeks or months, that aims to ensure continuity and quality of care and promote recovery at the interface between hospital and home, care home, primary care and community services [155]. In the Netherlands, the government implemented Short-Term Residential Care (STRC) in 2015 as part of major health care reforms because of rising expenses [108]. STRC is bed-based care for general health problems that do not require admission to the hospital, but also cannot be treated at home [171]. The goal of STRC is to enable older adults to return home and live independently in the community.

In the Netherlands, three types of STRC are provided: low-complex, high-complex and hospice care [123]. Low-complex care only provides care in Activities of Daily Living (ADL) and the general practitioner is the responsible physician. In addition to ADL care, STRC high-complex provides 1.5 hours of treatment per week such as physical therapy, medical diagnostics or occupational therapy. STRC hospice care provides care in the last three months of life.

4.2.2 Flow diagram

The admission control for all STRC locations in The Netherlands is regionally centralized. For that reason, we use one model at the aggregated level of a complete region. Figure 4.1 shows a flow diagram of the patient journeys through STRC.

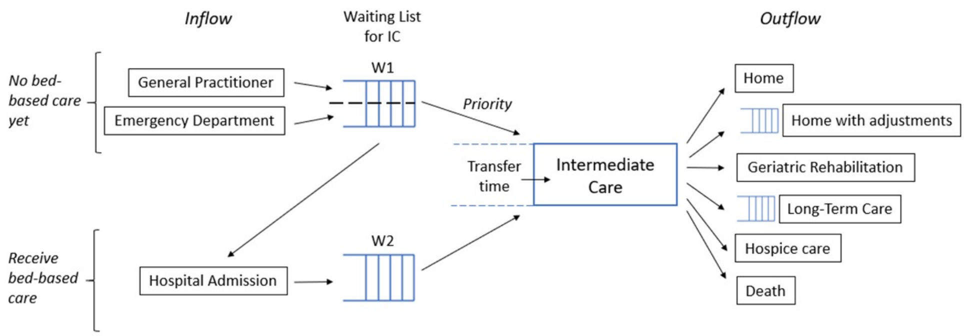


Figure 4.1: Current practice flow diagram STRC

In Figure 4.1, we see that admission to STRC occurs via two waiting lists. The first waiting list (W1) is for admissions from the home situation. The patient is referred by the General Practitioner (GP) or Emergency Department (ED) and is not in need of hospital-level care. The second waiting list (W2) is for patients residing in the hospital (HA) needing extra time to recover in intermediate care. The patients at W1 have priority over the patients at W2, since W1-patients often have a pressing need for care that cannot be provided at home. Moreover, patients from the ED need to be admitted to STRC within the same day, because otherwise a hospital admission is required. After the stay in STRC, the patient is discharged to home, is discharged to another care provider, or has died.

Furthermore, because of staff shortages, STRC admissions in the evenings or weekends are often not possible. For the weekends we have admission data that confirms this, see Figure 4.2. Finally, a transfer time is needed to do the administration and transfer of an older adult to the STRC location. By expert discussion, this transfer time was estimated to be on average 1.5 days.

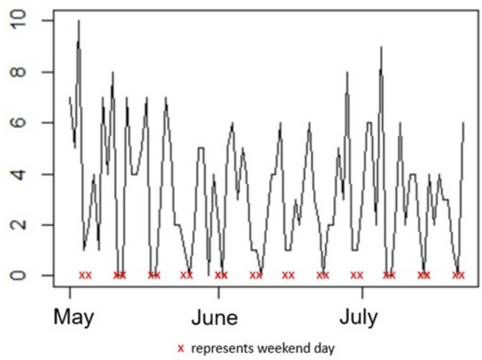


Figure 4.2: Number of STRC admissions in 2019 showing a decrease during weekend days.

4.2.3 Data collection

Various data sources containing STRC admissions were collected. In Amsterdam, almost all STRC care is high-complex, which is therefore the focus of this research. Henceforth, the term STRC will be used for our focus care type: high-complex STRC. We only used data from 2019, such that COVID-19 did not affect the patient flows. The following data sources were used:

- Non-public microdata of Statistics Netherlands (SN) [39]. This data source contained health care expenses at the individual level from which bed occupation, admissions, Length-of-Stay (LoS) and discharge destinations could be determined.
- Hospital STRC. This data was from an STRC that is located in a hospital. This STRC functions as a real-life case study for what happens if the admission to STRC is not hindered by logistic inefficiencies.
- Data from the centralized admission portal. This data source provided us with the number of STRC beds in Amsterdam and the GP waiting times for STRC.

We needed to receive formal permission to make use of the microdata database, and this data could only be accessed via a secured portal. Both other datasets were anonymized and did not contain sensitive patient data. Concerning the ethics: this study does not fall under the scope of the Medical Research Subjects Act [41].

4.2.4 Alternative policy scenarios

Current organizational problems of the STRC were collected in the study of the previous chapter [171]. In addition, we consulted employees of the centralized admission portal and healthcare organizations facilitating STRC. The problems that were deemed most urgent were: lack of possibilities to admit patients during the evening and weekends, missing information during the admission process, shortage of time and money during an admission to treat patients and lacking adjustments at the home of the older adult admitted to STRC. These problems were translated into possible process improvements, which are evaluated in the simulation as scenarios. We distinguish the alternatives (1) *in front*, (2) *during admission*, and (3) *at the back*. The following scenarios were obtained.

Alternatives in front

Admission Turns (AT). The proposed alternative is to compensate STRC locations for keeping beds available in the evening for admission by turn. This means that every other evening/night/weekend day a specific location is responsible for admissions from the GP, ED or hospital. We model this by allowing 24/7 STRC admissions and a transfer time of four hours.

Triage Ward (TW). Another scenario is to set up a triage ward, as illustrated in Figure 2. The triage ward is intended for older patients for which it is unclear what type of care is needed. The maximum LoS in the triage ward is fourteen days, after which patients are discharged to STRC or another destination. We model the TW by allowing 24/7 STRC admissions and a transfer time from the hospital to the TW of four hours. The average transfer time from the TW to the STRC remains 1.5 days. For a flow diagram of this scenario, see Figure 4.3. Note here that the inflow refers to the same inflow process as in Figure 4.1, but now the inflow is first accessing the TW prior to possibly entering IC.

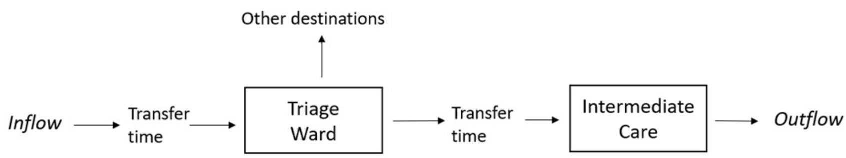


Figure 4.3: Scenario Triage Ward.

Alternative during admission

Higher Tariff (HT). The current tariff that STRC organizations receive from health insurers is to provide up to 1.5 hours of diagnostics and treatment per week, which is rather low compared to the high intensity of needed care [171]. The proposed alternative is to increase the tariff, such that more hours of diagnostics and treatment are possible and multiple treatment types can be provided in parallel. As a result, the LoS becomes shorter.

Alternative at the back

Reduced Bed-Blocking (RBB). Patient discharges can be delayed due to the lack of timely home adjustments required for the patient to return home. Such home adjustments are for example a stairlift or supporting handles in the shower. In addition, waiting lines for Long-Term Care (LTC) can be shortened by increased facility capacity or by improved waiting-line management [10]. We model the effect of this policy by a reduced LoS for the patients that require home adjustments or need LTC.

4.2.5 Simulation

After the process map was obtained and the data was collected, the DES model was developed. DES is a mathematical technique in which process characteristics are identified and modeled as a sequence of events in time. Examples of events are the arrival of a new patient and a patient's discharge. Using the DES, the impact of alternative policies can be evaluated. The method DES was the preferred method for this study as it is a stochastic method capturing individual variability, and thus suitable for estimating waiting times - which is a key aspect of interest here. More information about the DES method can be found in the literature [162].

The simulation was built from scratch in Python following the batch mean method. The warming-up period consisted of 5000 client departures. This number was found to be sufficient by observing that the mean waiting times were not consistently increasing after the warming-up period. The batch length consisted of 1000 client departures and the number of subruns was set to 100 for each scenario.

Using the data, we checked which probability distributions could be used best for modeling the arrivals and LoS. The arrival process is modeled as a Poisson process in which we included some time restrictions. For the ED arrivals, we use evidence from the literature that most ED patients arrive between 8AM and 12PM [115]. To perform the simulation, we need to identify the probability distributions underlying the arrival processes and LoS.

Arrivals
For the distribution of the arrivals, we plotted the weekly number of arrivals to IC and the Poisson distribution that was fit on the data, see Figure 4.4.

Moreover, a chi-square test indicated that the number of arrivals per week can be according to a Poisson distribution $\chi^2(df = 15, N = 52) = 11.0, p > .05$.

Length-of-Stay
The LoS was identified per outflow type and plotted together with the best fitted probability distribution, see Figure 4.5.

We tested statistically whether the data could be the realization of the underlying probability distributions. For this, we used the Kolmogorov-Smirnov tests, with as null-hypothesis that

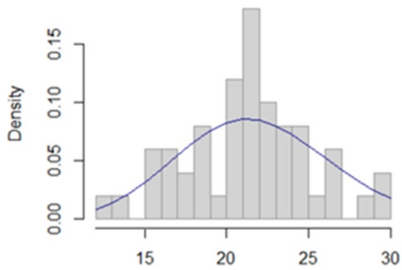
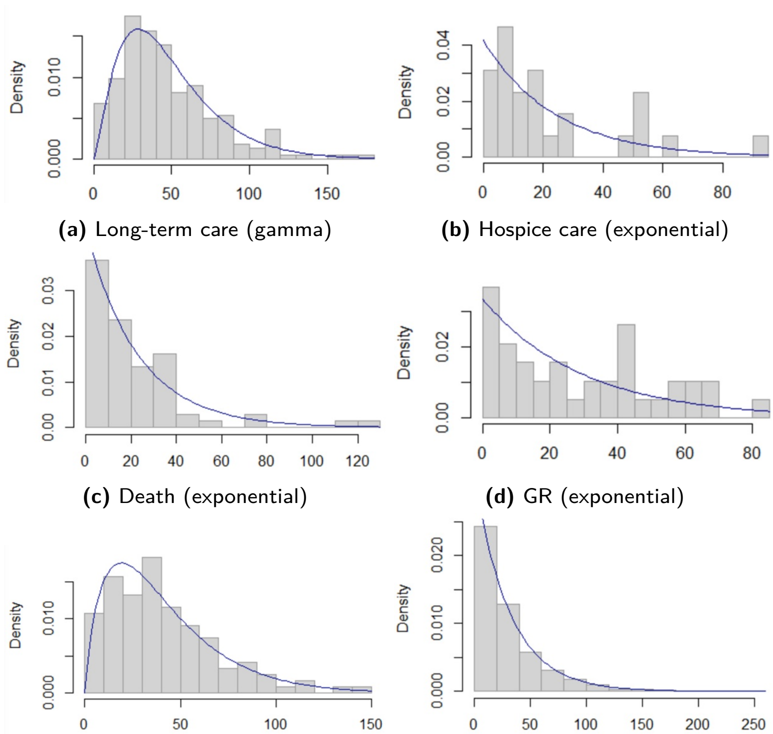


Figure 4.4: Density plot with weekly arrivals (bars) and fitted Poisson distribution (line). Note that the Poisson distribution is only defined at integer values.



(a) Long-term care (gamma)
(b) Hospice care (exponential)
(c) Death (exponential)
(d) GR (exponential)
(e) Home with adjustments (gamma)
(f) Home (exponential)

Figure 4.5: Density plot per outflow type with lengths-of-stay (bars) and best fitted distribution (line).

the empirical distributions are sample realizations of the proposed distributions. For all distributions, the null-hypotheses were not rejected at a 95% confidence level ($p > .05$).

Model validation
To validate our model, i.e., to check if our current practice simulation represents the real situation, we compared the distribution of occupied beds resulting from our simulation with real data based on non-public data from SN. Figure 4.6 shows that the simulated occupied bed distribution is quite similar to the real data. The mean occupancy is 84% (simulated) versus 85% (real data). The figure also showed that the real data is slightly more concentrated around the mean than the simulated data. Figure 4.7 shows boxplots of the simulated waiting times until placement in STRC next to the real waiting times data. The results show that the boxplots are rather similar except for the lowest quarter. An explanation for this is that the waiting times in the real data are only on day-basis and not per hour, which leads to 25% of the waiting times to be exactly zero.

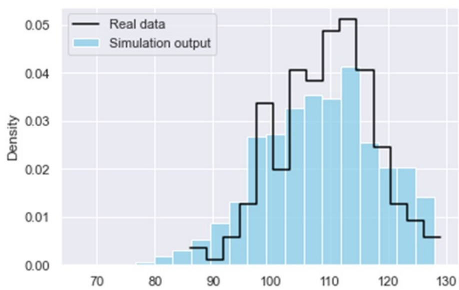


Figure 4.6: Occupied bed distribution

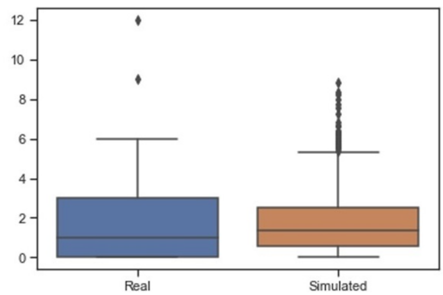


Figure 4.7: Waiting times from GP

The outcome measures that we aim to evaluate in the DES are presented and elucidated in Table 4.1.

4.3 Results

Now, the results are discussed of both the data analysis and the simulation.

4.3.1 Flow data

It should be noted that with the data only realized care journeys could be determined that might differ from the intended ones because of logistic bottlenecks. For STRC, the inflows might be affected by hospital admissions of patients for which timely STRC admission was not possible. We find in the data of the Hospital STRC in 2019 that 117 patients were transferred from the ED, and 131 patients were admitted after a hospital admission. This

Outcome measure	Description
Waiting time hospital	The average time between a planned hospital discharge and placement in STRC
Waiting time GP	The average time between STRC application by the GP and placement in STRC
Percentage of non-urgent (N-U) patients placed within 3 days	The percentage of older adults that are non-urgent, ie patients referred by the GP or from the hospital, for which the time between application and placement in STRC is less than 3 days
Percentage of patients presented at the ED that need a hospital admission	Percentage of older adults at the ED for which no STRC admission was possible within the same day, for which a hospital admission was necessary
Number of patient replacements	The average number of older adult replacements, where we count admission to STRC, hospital admission from the ED, admission to the Triage Ward, and STRC discharge as replacements
Number of patients delayed per day	The average number of older adults per day that are waiting in a hospital bed for an STRC bed

Table 4.1: Description of outcome measures.

implies that out of the patients who arrived via a care form, the inflow of older adults from the ED was 47% and from a hospital admission was 53%. Moreover, from the microdata of SN we also obtain inflow percentages for Amsterdam, of which 25% of the older adults transfer from the ED and 75% from a hospital admission. This large deviation of inflow percentages from the Hospital STRC might indicate that circa half of the current ED visits leads to a redundant hospital admission.

The inflow values of the ED and Hospital Admission were identified by the microdata source, which led to a total of 12.4 older adults' admissions per week. These were distributed according to the 47:53 ratio as found in the data, which led to 5.8 and 6.6 admissions per week. The results of the data analysis are provided in Table 4.2. It can be seen that the LoS for Home Care (HC) with adjustments and LTC is the longest. The expected reason for this is that patients have to wait before these services are available before they can be transferred. However, it can also be the case that these patients have more complex problems, which results in a longer reablement.

4.3.2 Number of beds

Now that we have validated the current practice model in section 4.2.5, we can evaluate the results of other scenarios. First, we check if the waiting times are the result of a lack in number of beds. For that purpose, we run the current practice simulation with varying number of beds. The results are shown in Figure 4.8.

Figure 4.8 shows that, most surprisingly, the current number of beds (128) is not the cause of the waiting times, since increasing the number of beds does not lead to a significant reduction. Moreover, it becomes clear that the patients from the GP have priority over

Parameter	Mean value
Inflow requests	
General Practitioner	9.4 older adults/week
Emergency Department	5.8 older adults/week
Hospital Admission	6.6 older adults/week
Outflow fraction	
Home	57.8%
Home with adjustments	10.7%
Long-Term Care	19.8%
Geriatric Rehabilitation	3.4%
Hospice Care	2.3%
Death	6.0%
Average length-of-stay	
STRC Home	31.1 days
STRC Home with adjustments	43.9 days
STRC Long-Term Care	47.8 days
STRC Geriatric Rehabilitation	29.8 days
STRC Hospice Care	22.9 days
STRC Death	22.9 days
Hospital Admission	5.2 days
Number of beds STRC	128

Table 4.2: Flow data describing the inflow, outflow and average length-of-stay of the STRC facilities.

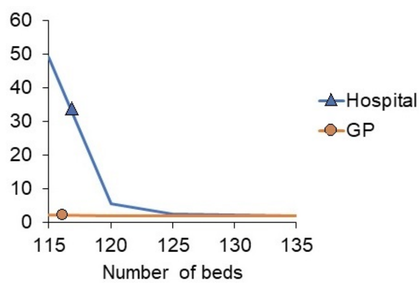


Figure 4.8: Expected waiting time in days for STRC for varying numbers of beds.

patients from the hospital. The average waiting time of 1.8 days is thus the result of the transfer time (which is on average 1.5 days) and the result of the lack of admission possibilities outside office hours. Alternative policies might reduce the waiting times, which we explore in the next step.

4.3.3 Alternative policies

The current practice and alternative policies were simulated for varying numbers of beds, which resulted in the output provided in Table 4.3. Confidence intervals of simulation outputs provide the statistical bounds on the means [98]. Our simulations were run until the width of the 95% confidence intervals on the waiting time for the hospital was less than 0.1 days. Therefore, the results were presented as point estimates without the intervals. The descriptions of the outcome measures can be found in Table 4.1.

We see that in the current practice scenario the fraction of ED visits leading to a hospital admission is 51%, which is in line with the expectation of the current process. We also see that only with a policy aiming at the admission access (i.e. AT and TW) the waiting times can be reduced. This is in line with our earlier observation that the waiting times are a result of an inefficient admission process. Moreover, for these policies, the hospital admissions from the emergency department are reduced. We see that the TW performs equally to all outcome measures compared to the AT policy, except for the number of patient replacements which increases from 2.0 to 2.7.

With the HT scenario, we have that if the LoS can be reduced by 30%, the number of beds can be reduced from 128 to 100 without an increase in waiting time. The bed-blocking scenario also leads to possible reductions in number of beds. The gain is the largest when reducing the LTC for the patients that transfer to LTC, since these current LTC are longer and the fraction of patients is the highest.

Finally, some combinations of alternative policies were run. We see that if both the AT and HT policies are implemented, the number of STRC beds can be reduced to 100, while also the waiting times are reduced, the number of hospital admissions is reduced and the number of patients in the wrong beds are almost zero.

4.4 Discussion

In this research, a simulation study was performed to evaluate the effects of alternative policies to improve intermediate care. For a case study of the city of Amsterdam, The Netherlands, it was found that the current waiting times for the intermediate care facilities are not due to a shortage in the number of beds, but are the result of an inefficient admission process. We show that if the organizations make the application process more efficient and extend the office hours' time frame with an evening and weekend admission policy, the number of unwanted hospital admissions prior to intermediate care can be decreased from approximately 51% to 3%.

Scenario	Number of beds	Waiting time hospital	Waiting time GP	N-U placed < 3 days	% ED → hospital admission	Nr. of repacements	Nr. of patients delayed / day
Current practice	128	2.4 d	1.8 d	79%	51%	2.1	3.3
Admission Turns							
	128	0.2 d	0.1 d	100%	3%	2.0	0.2
	115	3.8 d	0.3 d	88%	33%	2.1	4.6
Triage Ward							
TW = 48	80	0.2 d	0.1 d	100%	3%	2.7	0.2
TW = 40	80	3.8 d	0.6 d	85%	30%	2.8	8.1
Higher Tariff							
LoS reduction of 30%	128	1.8 d	1.8 d	82%	46%	2.1	2.4
LoS reduction of 30%	100	1.9 d	1.8 d	82%	48%	2.1	2.5
LoS reduction of 30%	100	3.3 d	1.9 d	75%	58%	2.2	4.7
Reduce Bed-Blocking							
HWA LoS reduction of 30%	128	1.9 d	1.8 d	82%	48%	2.1	2.6
LTC LoS reduction of 30%	128	1.8 d	1.8 d	82%	47%	2.1	2.4
LTC LoS reduction of 30%	110	8.0 d	2.0 d	66%	67%	2.2	12.0
Combinations							
AT + HT (LoS reduction of 30%)	100	0.1 d	0.1 d	100%	0%	2.0	0.1
AT + HT (LoS reduction of 30%)	80	6.3 d	0.4 d	83%	45%	2.1	8.2

Table 4.3: Simulation results: comparison between current practice and alternative scenarios.

In addition, if a higher tariff is introduced that reduces the LoS by 30%, the number of beds can be decreased by 22% without an increase in waiting times. A reduced LoS can also be obtained by cutting down waiting times for follow-up care, but since this concerns only a fraction of the patient outflow, the effects are smaller. The TW policy leads to the unwanted consequence that a high number of patient replacements need to be performed. This policy should therefore only be implemented if it supports the care process in other aspects such as succeeding in allocating patients directly to the right care form. Finally, combining the right policies leads to the largest gain in reduced waiting times and bed capacity.

To the best of our knowledge, our approach to model the IC facilities as a mid-chain care form in the healthcare system in which we also consider the admission policies is unique. In an earlier IC simulation study, no admission policy is mentioned and it is assumed that the older adults wait until their best-fitted care form is available, which might be less realistic in acute situations [38]. In addition, a literature review of modeling studies for IC showed that little modeling studies focus on outflow to community health services, comparable to IC, and do not often focus on the connection to the use of acute care [78]. Finally, one study did construct the care pathway of IC, but the researchers were not able to test alternative scenarios [92].

Besides the possible improvements in patient satisfaction and health outcomes, we also might save costs if the process flow is improved. In particular, the average tariff for three days hospital admission due to older adults' problems is €900 per day [124]. Multiplied by the number of hospital days saved per day for the AT policy (3.2 days saved), an estimated €2880 per day can be saved due to hospital avoidance. The IC literature is primarily focused on the cost-effectiveness of the supported discharge goal of IC and not on admission avoidance [36]. This paper thus also contributes in that aspect. Moreover, it was found that especially admission avoidance leads to cost-savings [36], which is in line with this study as well.

Our study has several strengths and limitations. A valuable contribution is that we consider the whole chain of admission and discharge in which IC is involved. Furthermore, our study focuses on cost-effectiveness as well as on patient relevant outcomes such as number of transfers. A limitation of our study is that we could not include process features in our simulation for which we lacked data. Firstly, a small fraction of the IC beds is dedicated to observation of older adults who had a delirium and are recovering from this delirium. However, we do not have data about the patient arrivals for these specific beds, which are therefore not considered. Secondly, acute patients from the ED can be surpassed by (not acute) hospital patients, since the hospital patients already claimed a bed in the IC facility. As we do not have information about the number of cases in which this happens, this feature was also left out.

4.5 Conclusions and Implications

In this study, we found that simulation is an effective tool for modelling IC facilities and determining policies for improvement. In our case study, we show that with the right policies, fewer beds (and hence personnel) are needed for IC in Amsterdam. The improved access to these facilities implies that more hospital admissions can be avoided and bed-blocking in the hospital can be reduced. Moreover, the newly available beds can be dedicated to other care forms for which a shortage exists, such as LTC. Hence, an efficiency improvement of one pivotal care facility leads to an improvement of the older adults' care system as a whole.

Currently, decisions on healthcare interventions and capacities are often made without sound quantitative insights into the effects of the decisions. This research shows the importance of a simulation approach for improvement of the care domain of older adults. This novel strategy can prevent high costs and effort put into multiple pilot studies of which it is unclear whether the intended outcomes will be obtained. For example, in Amsterdam a pilot study could have been set up to reduce the LoS of the IC facilities which would not have led to a decrease in waiting times. Therefore, a data-driven approach is essential to identify problems' true causes and appropriate solutions, which results in increased cost-effectiveness and patient satisfaction.

Part II

Long-Term Care

5

Allocation of Older Persons to Nursing Homes: Theory

Based on:

Arntzen, R. J., Bekker, R., & van der Mei, R. D. (2024b). Preference-based allocation of patients to nursing homes. *Operations Research for Health Care*, *42*, 100442.

Abstract

Objective: In many countries, the rapidly aging of the population leads to an additional burden on already stretched long-term care systems. This often manifests itself in excessive waiting times for long-term care centers, and in abandonments (i.e., patients passing away while they are waiting). Interestingly, in practice, long waiting times are not solely caused by a lack of available total capacity in the system, but also by systematic inefficiencies in the allocation of patients, each with their personal preferences and (in)flexibility, to geographically distributed care centers.

Method: Motivated by this, we propose a new and easy-to-implement method for the optimal allocation of patients-in-need to nursing homes, balancing the trade-off between the waiting time performance and the individual patients' preferences and levels of flexibility.

Results: The optimal placement policy found by solving a Markov Decision Process demonstrates that for small instances, the mean optimality gap of the allocation model is equal to 1.3%.

Conclusion: With the allocation model, individual preferences can be served better, which thus provides a powerful means to face the increasing need for patient-centered and sustainable long-term care solutions.

5.1 Introduction

Today, deficits in local nursing home capacity often lead to overcrowded waiting lists [63, 132], and the aging population causes longer waiting times in prospect. Lack of timely access to Long-Term Care (LTC) is a crucial issue due to the resulting confusion, distress, and anxiety of patients [102]. In addition, during the delay in admission, informal caregivers experience depressive symptoms and feelings of burden [117]. In addition, high costs are involved for the healthcare system because waiting patients occupy hospital beds [111], and patients may suffer injuries due to omission of proper care [16].

In addition to the costly solution of increasing facility capacity, waiting list management can be used to get patients to the right place more quickly. However, little research has been done on how long-term care waiting lists should be managed or organized [42, 138]. Even with this limited amount of research, almost all studies focus on improving available capacity (e.g., implementing the use of transitional care facilities [52]) and *not* on the *organization* of the waiting list itself. The Operational Research community is explicitly called upon to develop models for effective healthcare waiting lists [27]. Nevertheless, hardly any mathematical models have yet been developed that can be used to study this highly socially relevant issue.

A core characteristic that distinguishes LTC from other domains in terms of waiting list management is that *patients* have individual *preferences* for nursing homes. The possibility of including preferences appropriately in designing a waiting-list policy is currently lacking in relevant research fields such as the queueing literature on routing and admission control. The contribution of this chapter is two-fold. First, we propose a waiting-list management method that incorporates routing preferences to parallel facilities. Second, we provide a solution to the currently existing waiting list problem in nursing home care. More specifically, our research objective is to design an allocation model for waiting list management policies that is (i) easy to implement and understandable for healthcare employees, (ii) scalable as the number of patients and nursing home beds can be large, and (iii) able to combine the goals of retaining preferences best and keeping waiting times low.

5.2 Literature

LTC is provided for patients who have physical and/or mental deficits that prevent them from performing the usual daily tasks. The majority of LTC is utilized by the elderly (65+), and therefore most LTC facilities are designed for this sub-population. LTC can be divided into three categories: (1) *informal care*, (2) *community care*, and (3) *institutional care* [161]. Informal care is provided by the patient's family, friends, and neighbors and consists mainly of basic activities such as cooking a meal or walking outside. It has been found that 92% of patients who live at home receive voluntary help [93]. Community care involves paid services targeted at patients who live at home and consists of a broad range of services from nursing care at home to day care at an external location. If home is not an appropriate living environment for the patient's condition, institutional care is required. Institutional facilities in which specialized care is provided by nurses are referred to as typical 'nursing homes'.

Proper access to nursing homes cannot always be guaranteed due to excessive waiting times. The European Commission reports high proportions of older people in need of institutional care who are currently on waiting lists in the EU [63]. For example, these fractions of older people on the waiting list are 16% in the Netherlands, 33% in Bulgaria, and 53% in Hungary. The problem exists in countries all over the world; see, for example, [15, 68, 70] for some reports in Australia, Canada, and South Africa. The lack of regional nursing home capacity can further exacerbate the problem. For example, in Copenhagen, the waiting time for a bed in a nursing home, which is not necessarily a preferred one, is over 3.5 years. Even in countries where waiting lists for nursing homes are not (yet) an issue, we observe regional and popularity issues with certain nursing homes. In the USA, only 29% of all residential care communities reported having a waiting list, but among those, the average waiting time for admission was around five months [95].

In the following, we address the literature regarding the management of waiting lists for nursing homes from different angles: waiting for nursing homes and waiting for healthcare in general. Thereafter, we elaborate on the relation of our research with the literature on routing and conclude with our contribution.

5.2.1 Waiting-list management in long-term care

The LTC waiting process starts with the selection of the right nursing home to apply for, based mainly on patient preferences, such as location [158] and cultural factors [156]. Specific preferences, such as cultural background, significantly influence waiting times [156, p. 1348]. If a patient's application to a nursing home is not immediately accepted, the waiting period begins. To better manage the waiting period, interventions such as transition care settings and adjustment of home care levels are explored. For example, Crotty, Whitehead, Wundke, Giles, D. Ben-Tovim and Phillips [52] found that off-site transitional care units can be utilized during the waiting period without adversely affecting patients' conditions. However, some patients may prefer not to use such units, suggesting a voluntary approach. Furthermore, research by Pedlar and Walker [139] highlights the preference for increased home care over nursing home admission by 90% of patients on waiting lists. This preference can significantly reduce nursing home demand and lead to substantial cost savings, given the higher expense of nursing home care compared to home care.

Limited research exists on the management of waiting lists themselves. In a meta-analysis by Chafe, Coyte and Sears [42], only two papers on this topic were identified. The first paper by Burkell, Wright, Hoffmaster and Webb [34] examined the change from a First-Come-First-Served (FCFS) policy to one that prioritizes patients based on urgency. They observed significant variability in patient care needs, suggesting that a needs-based criterion would drastically alter priority. Similarly, in a study by Meiland, Danse, Wendte, Gunning-Schepers and Klazinga [116], priority was based on urgency, revealing that non-urgent patients experienced longer waits without deterioration of their condition. Moreover, their satisfaction levels remained consistent while waiting at home, indicating that prioritizing urgent patients had minimal negative consequences for non-urgent ones.

In [138], waiting-list management is discussed in a more general setting, including patients discharged from hospitals who need residential care. A Markov Decision Process (MDP) was developed to determine when patients in hospitals should receive priority for nursing home placement over those waiting at home. However, lowering the threshold simply shifts the waiting time problem to patients at home. Furthermore, a simulation tool estimated that nursing home stays must decrease by two to three years to meet service level targets.

5.2.2 Waiting-list management in healthcare

The impact of priority settings is also studied in other health contexts. Bowers [29] aimed to simulate the waiting lists of the UK's National Health Service. By studying empirical waiting-time distributions, it was found that a FCFS policy did not correctly describe the behavior of the waiting lists. Motivated by this, a model was developed that includes priority parameters to differentiate between urgent and routine patients and to include specified target waits for those different groups of patients. Although the goodness-of-fit test revealed that the developed model could not adequately describe empirical data, the model can be used to predict waiting times when such alternative policies are implemented.

The priority setting can depend on more factors than just urgency. However, determining the right priority instruments is complex. In 1996, the International Society for Priorities in Health was founded to investigate health priorities for both theoretical and practical purposes [89]. Priority settings may even be more complicated when demand is heterogeneous, such as in the case of surgical procedures for which the expected outcome depends on the characteristics of the patient. There are even settings in which both demand and supply are heterogeneous, such as in the allocation of kidneys to patients where each kidney-patient combination determines the expected lifetime of the patient, as seen in Zenios, Chertow and Wein [198]. In those cases, a regular waiting list will not suffice, and an allocation model is needed.

5.2.3 Routing models

The issue of waiting-list management can also be examined from a queueing-theoretic perspective. There is now a substantial body of literature on routing customers to parallel queues. For our purposes, it is most intriguing to consider queues with both heterogeneous customers and heterogeneous server pools. Such systems have traditionally been studied in the call center domain [67, 187], where the assignment of customers to server pools is known as Skill-Based Routing (SBR). As indicated in [44], the analysis of SBR is typically intractable. Therefore, optimal routing decisions often rely on asymptotic approximations (see, e.g., [74, 75, 109]) and special cases, such as the V, inverted V, N, and X designs. For additional background and references, we refer the reader to the recent survey in [44]. This survey also discusses the complexities of SBR models in healthcare applications, focusing on in-hospital patient flow from emergency departments to hospital wards. It should be noted that [44] also explores an MDP formulation for optimal routing. Due to the explosion of the state (and action) space, the authors only considered small-scale examples. For a recent study using approximate dynamic programming techniques for somewhat larger instances, see [54].

In computer-communication applications, the number of server pools N is generally considerably larger than in our nursing home setting. To avoid excessive communication overhead, power-of-d routing policies have been suggested. For such a policy, $1 \leq d(N) \leq N$ of the N server pools are randomly selected, and the customer is routed to the shortest of those queues. Even in the case of sampling two queues ($d(N) = 2$), waiting times are drastically reduced (see, e.g. [119]). In fact, [121], among others, demonstrate that for relatively small $d(N)$, asymptotically optimal behavior can be preserved. We note that these models fundamentally differ from our setting, as customers are routed to a single queue upon arrival, and there are no preferences (server pools are homogeneous). Nonetheless, the power-of-d results indicate the enormous potential to add some flexibility in the allocation of patients to nursing homes.

5.2.4 Our contribution

Long waiting lists for nursing homes are a significant societal problem, yet existing research often overlooks a crucial aspect: incorporating patient preferences into nursing home assignments. To fill this gap, we propose a method that integrates patient preferences into waiting-list management policies. Our aim is to develop a system that is easy for healthcare employees to implement and understand, that is scalable to accommodate varying numbers of patients and nursing homes, and that effectively balances patient preference retention with minimizing waiting times.

In addition to individual nursing home preferences, we also consider patients' preferences regarding willingness to wait. Drawing from the concept of utilities, as studied by Nomura, Yamori, Takahashi, Miyoshi and Tanaka [128] in content delivery networks, we formalize this aspect. Using utilities as a measure of patient preferences, our model offers a practical and comprehensible approach for healthcare managers. This patient-centered allocation method not only aligns with the individuality of patients but also allows efficient bed allocation by leveraging explicit preferences.

5.3 Context and Problem Description

In this section, we define the scope of the problem, which involves the system that incorporates the placement of patients in nursing homes.

5.3.1 Background and terminology

For the sake of clarity, we elucidate the terms used throughout the chapter.

Patients: We consider patients on a waiting list for nursing home placement, excluding high-emergency cases. All patients require the same type of care, allowing nursing home beds to be used interchangeably. Patients enter the system upon expressing readiness for placement, but remain at home while waiting.

Nursing home capacity: According to Ribbe [143], nursing homes are institutions that provide nursing, psycho-social, and personal care, as well as room and board. The capacity of a nursing home location is expressed in terms of the number of beds available at the location. The available number of beds may either be determined by the physical beds present at the nursing home or may depend on the availability of other resources such as personnel and services.

Preferences: In practice, certain nursing homes may be more appealing to patients due to factors such as proximity to their home, familiarity with the coordinating care organization, and specialization in specific subpopulations (e.g., LGBTQ+ patients or patients with the same ethnic background [156]). We assume that each patient has at least one preferred nursing home where they would ideally like to reside permanently. If a patient is placed in a nursing home other than one of their *preferred* choices, the placement may be temporary.

5.3.2 Problem outline

The conventional procedure involves patients enrolling in the waiting list(s) of their preferred nursing home(s) upon arrival. However, we propose an alternative that allocates patients to nursing homes to satisfy as many individual preferences as possible. Let $N = 0, 1, \ldots, n_{\max}$ denote the set of nursing homes, where we define *nursing home 0* as the location of the patient's home. We denote by b_n the number of beds available in the nursing home n, and let $b = \sum_{n \in N \setminus 0} b_n$ be the total number of beds. Moreover, let P be the set of patients who need a bed in a nursing home. Each patient $p \in P$ has a set of preferred nursing homes, $L_p \subseteq N \setminus 0$, in which they would like to reside permanently. If a patient is placed in a nursing home in the set $M_p := N \setminus L_p$, the placement is temporary and the patient still waits for a bed in a preferred nursing home.

Patients are assumed to arrive according to a Poisson process with rate λ. The service times in nursing homes are assumed to be exponentially distributed with mean $1/\mu$. Finally, the time until abandonment from the queue, resulting from patients dying or moving to another region, is exponentially distributed with mean $1/\theta$.

The patient placement process unfolds over time and space, as shown in Figure 5.1. Initially, the patient joins the waiting list for their preferred nursing home while remaining at home. As time elapses and the patient's condition worsens, a preference for temporary placement in a nursing home may arise. If a bed is available in a temporary facility, the patient can be relocated there while still awaiting placement at their preferred home. Upon availability, the patient moves to the preferred nursing home, staying until they leave the system, often due to death or transfer. Each patient's journey is unique, with factors like disease affecting progression.

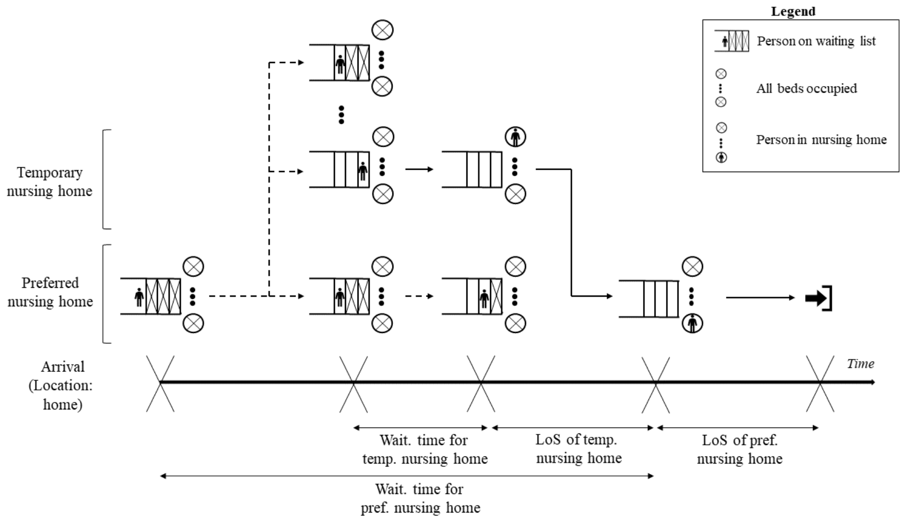


Figure 5.1: Example patient journey.

Objective and performance measures

The goal is to efficiently allocate patients to nursing homes while incorporating their preferences. This involves optimizing the aggregate utility of patients based on preferences for nursing homes and willingness-to-wait. We evaluated the trade-off between adhering to preferences and efficiency using the key metrics detailed in Table 5.1. Each patient p is assigned a preference, expressed in terms of g_{pn}, for each nursing home n, the preferred nursing homes being those with the highest values. Negative values ($g_{pn} < 0$) indicate a patient's unwillingness to be placed in a specific nursing home. We also account for patients' willingness to wait, formalizing this trade-off through utility functions; see Subsection 5.4.1 for a further discussion with an emphasis on the waiting aspect.

We allow each patient to have unique utility structures, offering customization appreciated by healthcare professionals. However, in practice, employing a limited set of patient profiles can streamline operations. These profiles capture common utility structures, simplifying the acquisition of patient preferences. Our model presents individual utility structures, while practical examples in Section 5.5 illustrate the use of patient profiles.

Performance measure	Description
Waiting time	Time between arrival to waiting list and first and/or final placement
Abandonments	Fraction of patients that abandon the system
% preferred NH	Fraction of patients that end up in their preferred nursing home
Replacements	Average number of replacements per patient

Table 5.1: Performance measures.

5.4 Method

We first formalize the preference profiles and willingness-to-wait using utility functions in Section 5.4.1. These utility functions are used in our allocation model for assigning patients to nursing homes in Section 5.4.2. In Section 5.4.3, we evaluate the allocation model using simulation. We discuss some benchmark policies in Section 5.4.4.

5.4.1 Utility functions

We established the utility functions in collaboration with and with approval of experts in the elderly care domain. Specifically, the application of utility functions is an intuitive and attractive way for health professionals to reflect the impact of preferences for nursing homes and waiting on the choices of individual patients. In this setting, the utility for each nursing home $n \in N$ represents the 'level of happiness' to reside there.

The utility depends on the initial preferences of the patient p, the current location l_p and the elapsed waiting time w_p, $\forall\, p \in P$. We denote the utility function for patient p toward nursing home n as $u_{pn}(l_p, w_p), \forall\, p \in P, n \in N$. In particular, it should comprise the following four elements (with corresponding notation):

(i) Each patient $p \in P$ can indicate initial preferences for each nursing home $n \in N$. These preferences are denoted by g_{pn}.

(ii) The patient's $p \in P$ willingness to be placed in a temporary nursing home $n \in M_p \setminus 0$ increases with the waiting time w_p. This is described by the function $v^U_{pn}(w_p)$, which is strictly increasing in w_p.

(iii) The patient's $p \in P$ willingness to be placed in preferred nursing home $n \in L_p$ increases with the waiting time w_p. This is described by the function $v^F_{pn}(w_p)$, which is strictly increasing in w_p.

(iv) Patient's $p \in P$ relocation between two temporary nursing homes results in a utility loss. This utility loss is modeled by the replacement penalty K.

The preferences (i) remain constant over time, but if a patient's preferences change during the waiting period due to new information about nursing home features, they can update their preferences. The impact of waiting time is captured by elements (ii) and (iii). Specifically, $v^U_{pn}(w_p)$ reflects the flexibility of a patient in waiting time. Patients who prefer a quicker placement in a temporary nursing home will see their utility increase more rapidly with waiting, ensuring that they are placed sooner. However, this faster placement comes at the expense of their opportunity for rapid placement in a preferred nursing home, as reflected in a less steep increase in $v^F_{pn}(w_p)$ for preferred nursing homes. The waiting time utility function $v^U_{pn}(w_p)$ is added to the utility of temporary nursing homes if a patient resides at home. Thus, a patient may initially have a negative utility for a temporary nursing home, but as the waiting time increases, their utility may become positive, leading them to prefer the temporary nursing home over staying at home.

Regarding element (iv), patient relocations between temporary nursing homes are only considered if the increase in utility outweighs a replacement penalty K. Patients typically only relocate between temporary nursing homes if the utility gain exceeds this penalty. However, if a patient's preference is a preferred nursing home, the replacement penalty is not included.

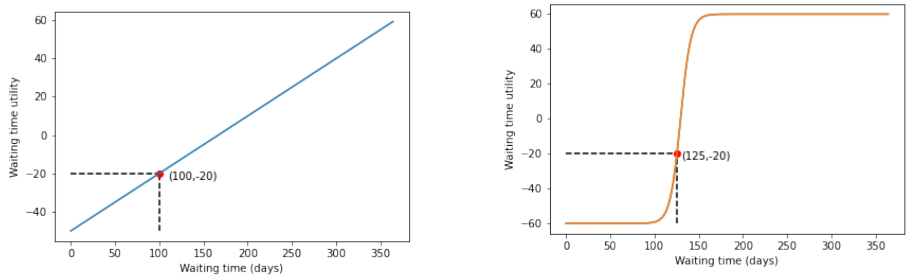
Waiting-time utility

The waiting-time utility functions can be chosen so that the preferences of individual patients are best represented. In the following, we discuss two different forms of utility functions. First, a simple choice for waiting-time utility is a linear function of the elapsed waiting time w_p of the form $v^U_{pn}(w_p) = \alpha_p w_p - \beta_p$, with $\alpha_p > 0,\ \beta_p \geq 0,\ p \in P$. In this case, the increase in waiting-time utility always grows at the same speed. However, according to elderly care experts, the increase in waiting-time utility is not constant. They describe the willingness to be placed in a nursing home as follows: Initially, the waiting time utility increases slowly, whereas after some 'breaking point' (e.g. an abrupt loss in functionality) the utility increases drastically, after which the utility increases slowly again. For that reason, a second more natural choice for the waiting time utility function is to use a sigmoidal function, thus of the form

$$v^U_{pn}(w_p) = \frac{\alpha_p}{1 + e^{-(\epsilon_p w_p - \gamma_p)}} - \beta_p \qquad \forall\ p \in P.$$

The parameter values of the utility functions determine the speed with which a person is willing to be placed in a nursing home. Patient $p \in P$ wants to be placed in nursing home $n \in N$ if $u_{pn}(l_p, w_p) > 0$, thus in the linear case if $g_{pn} + \alpha_p w_p - \beta_p > 0$. If $\beta_p > 0$, only if w_p is large enough we have $g_{pn} + \alpha_p w_p - \beta_p > 0$. An example of this for both waiting time utility functions is provided in Figure 5.2 below.

In Figure 5.2, we assume that $g_{pn} = 20$ for patient $p \in P$ and nursing home $n \in N$. Then, only if $v^U_{pn}(w_p) > -20$ we have $u(l_p, w_p) > 0$, after which the patient prefers to be placed in nursing home $n \in N$ over staying at home. In the linear case this holds if $w_p > 100$ and in the sigmoidal case if $w_p > 125$.



(a) Linear: $v^U_{pn}(w_p) = 0.3w_p - 50$

(b) Sigmoidal: $v^U_{pn}(w_p) = \frac{120}{1+e^{-(0.17w_p-22)}} - 60$

Figure 5.2: Example waiting time utility functions.

Now, the utility function of patient $p \in P$ at location $l_p \in M_p$ towards location $n \in N$ with elapsed waiting time w_p is defined as follows:

$$u_{pn}(l_p, w_p) = \begin{cases} 0 & \text{if } n := l_p, \\ g_{pn} + v^U_{pn}(w_p) & \text{if } l_p = 0,\ n \in M_p \setminus \{0\}, \\ g_{pn} - g_{pl_p} + v^F_{pn}(w_p) & \text{if } n \in L_p, \\ g_{pn} - g_{pl_p} - K & \text{if } l_p \neq 0,\ n \in M_p \setminus \{l_p\}, \\ -\infty & \text{if } n = 0,\ l_p \in M_p \setminus \{0\}. \end{cases} \tag{5.1}$$

In words, the utility is 0 if patient $p \in P$ stays at the same location. The utility is $g_{pn} + v^U_{pn}(w_p)$ if the patient lives at home and goes to a temporary nursing home. The utility is $g_{pn} - g_{pl_p} + v^F_{pn}(w_p)$ if the patient goes to a preferred nursing home. Furthermore, the utility is $g_{pn} - g_{pl_p} - K$ if a patient resides in a temporary nursing home and goes to another temporary nursing home. Finally, if the patient no longer lives at home, i.e., $l_p \in M_p \setminus 0$, the utility for home is set to $-\infty$, which prohibits a return placement. Adding the waiting-time utilities in this way ensures that if there are two patients $p', p'' \in P$ with the same utility functions and where $l_{p'} = l_{p''}$, but $w_{p'} > w_{p''}$, then patient p' is given priority over patient p'' for nursing home n. This fosters fairness in the system.

5.4.2 Allocation model

We propose a Binary Integer Program (BIP) model to allocate patients to nursing homes. The allocation model is based on a static setting (*snapshot*) of the dynamic process. The model solves the optimal allocation for the static setting so that the utility sum of all patients is maximized. For our allocation model, we are not interested in the patients who are already residing in a preferred nursing home, since these patients do not have to be re-allocated to another nursing home. Thus, we have that for each patient $p \in P$, the current location is $l_p \in M_p$. Note that the set M_p also includes the home location 0.

Similarly, we define the capacity c_n of nursing home $n \in N$ as the number of beds not assigned to patients $p \in P$ for whom $n \in L_p$. This is based on the reasoning that these patients already have a bed allocated. The capacity c_n thus includes beds that could be occupied by patients who temporarily reside there. Utility functions ensure that temporary patients are primarily allocated to the nursing home where they already reside, allowing them to stay in the same bed.

For the allocation problem, we define the decision variable x_{pn}, which equals 1 if patient $p \in P$ is placed at location $n \in N$, and 0 else. An overview of the notation is provided in Table 5.2.

Sets	
P	Patients
N	Nursing homes, with $n = 0$ as the home location
$L_p \subseteq N \setminus 0$	Preferred nursing homes of patient $p \in P$
M_p	Non-preferred nursing homes, i.e., $N \setminus L_p$
Parameters	
$u_{pn}(l_p, w_p)$	The utility of patient $p \in P$ at location l_p towards location $n \in N$, after waiting time w_p
l_p	The current location of patient $p \in P$, where $l_p \in M_p$
g_{pn}	The fixed utility of patient $p \in P$ for location $n \in N$
$v^U_{pn}(w_p)$	The waiting time utility of patient $p \in P$ until placed in nursing home after waiting time w_p
$v^F_{pn}(w_p)$	The waiting time utility of patient $p \in P$ until placed in permanent nursing home after waiting time w_p
K	The replacement penalty
b_n	The capacity of nursing home $n \in N$, where $b_0 = \infty$
c_n	The remaining capacity of nursing home $n \in N$, i.e., the number of beds not occupied by patients $p \in P$
Decision variables	
x_{pn}	Binary variable to indicate if patient $p \in P$ is placed at location $n \in N$

Table 5.2: Notation for the allocation model.

Now, the allocation problem can be formulated as follows:

$$\max \sum_{p \in P} \sum_{n \in N} u_{pn}(l_p, w_p) x_{pn} \tag{5.2}$$

$$\text{s.t.} \sum_{p \in P} x_{pn} \leq c_n \qquad \forall n \in N, \tag{5.3}$$

$$\sum_{n \in N} x_{pn} = 1 \qquad \forall p \in P, \tag{5.4}$$

$$x_{pn} \in \{0, 1\} \qquad \forall p \in P, n \in N. \tag{5.5}$$

The objective of the allocation problem (5.2) is to maximize the total utility of all patients $p \in P$. This must be done under the constraints that (5.3) no more patients can be allocated to a nursing home than the capacity allows and that (5.4) each patient can only be allocated to one location. Constraints (5.5) are binary constraints.

Since the constraint matrix is totally unimodular and the elements on the right-hand side of the constraints are integral, the solution of the linear problem is integral. Therefore, the allocation model can be solved using a linear solver, which makes it fast and scalable. Furthermore, the (binary) allocation problem is a special case of the Generalized Assignment Problem, in which the capacity needed for each job is equal to 1 [62].

Moreover, our waiting-time utility structure not only caters to patient preferences but also promotes efficiency. If a patient faces prolonged waiting times for a specific nursing home, their utility for alternative options may increase, causing placement in another facility. This design ensures optimal bed utilization, balancing patient preferences with efficient resource allocation.

5.4.3 Simulation model

In this section, we describe how the allocation model performs in a dynamic setting and can be evaluated using simulation. First, in line with the current procedure, we assume that patients can only be admitted to a nursing home at fixed regular moments (e.g. every morning). We define Δ as the time between two consecutive allocation moments. Then, the (deterministic) sequence of allocation moments can be described as $t_n = t_{n-1} + \Delta$, for $n = 1, 2, \ldots$. For example, if patients can enter once a day, Δ equals one day. The set of all allocation moments is denoted by $T = \{t_n\}_{n=1}^{\infty}$.

The time interval between the decision moments, Δ, determines the speed with which entry takes place. If $\Delta \to 0$, then entering a nursing home can be done instantaneously, such that the nursing home n behaves as an M/M/b_n+M queue. If Δ is chosen to be rather large, a bed might remain empty for some time, resulting in inefficiency in capacity use. On the other hand, a larger Δ implies that during a time interval more arrivals and departures take place, which increases the allocation possibilities.

More specifically, the model above with entering possibilities can be related to a queueing model with discharges at inspection instances [43]. In case of a single nursing home n and infinite patience (i.e., the M/M/b_n variant), the stability condition is

$$b_n \left(1 - e^{-\mu\Delta}\right) > \lambda\Delta,$$

where the term on the left-hand side corresponds to the maximum number of departures per Δ time units, which follows a Binomial(b_n, $(1 - e^{-\mu\Delta})$) distribution. Intuitively, $1 - e^{-\mu\Delta}$ can be interpreted as the effective service rate. Observe that $1 - e^{-\mu\Delta} \approx \mu\Delta$ for $\mu\Delta$ small, such that the capacity lost due to entering possibilities is small.

Now, in order to evaluate the allocation model in the dynamic setting, we implemented a simulation model, where the allocation model of Subsection 5.4.2 is executed based on a rolling horizon, i.e., at all instants t_n after updating arrivals, departures, and utilities. The procedure consists, after the initialization, of two steps: (1) incorporating the dynamic features of the simulation, and (2) running the allocation model. The two-step procedure is iteratively run until a stopping criterion is met. We refer to Appendix 5.A.1 for a more detailed description of the simulation procedure.

5.4.4 Benchmark policies

As mentioned, the aim of the allocation model is to have the 'best of both worlds', i.e., short waiting times, few abandonments, and placement in a preferred nursing home. Clearly, there are extreme policies that are best for one of such performance measures. Such policies provide insight into the best value that can be achieved, i.e., serve as a lower or upper bound. Ideally, the allocation model yields a performance close to those bounds.

Shared queue. This refers to the situation in which there is a single waiting list for all nursing homes (and preferences are completely neglected). The system is similar to the Erlang-C or Erlang-A model, with the capacity equal to the total number of beds of all nursing homes. This corresponds to the most efficient system design. Hence, the corresponding mean waiting times and fraction of abandonments provide lower bounds for the actual performance.

Separate queues. This corresponds to the other extreme, namely dedicated waiting lists for each individual nursing home. The queue of a single nursing home is similar to the Erlang-C or Erlang-A model, with capacity equal to the number of beds of the corresponding nursing home. Clearly, the number of preferred placements is now maximized (provides an upper bound), at the cost of waiting times and abandonments.

Optimal policy. An optimal policy can be formulated using an MDP that considers both direct and future costs in terms of preferred placements, abandonments, and waiting. For any reasonably sized instance, solving such an MDP is computationally prohibitive due to the curse of dimensionality. The goal of the optimal policy is to verify the optimality gap of the allocation model for very small instances. The precise MDP formulation and the size of the state space can be found in Appendix 5.A.2.

5.5 Results

In this section, we present numerical results from applying the Allocation Model (AM) to a small-scale setting. The aim is to gain insights into the model's performance and behavior under various parameter settings. Since our model facilitates in serving individual needs, we want to show the benefits of the allocation model for individual patients. For that purpose, selecting the same utility function for all patients will not suffice. Therefore, we created two characteristic groups of patients with similar utilities. The two utility groups are "Fast placement" (Fast Placement (FP)) and "Preferred placement" (Preferred Placement (PP)). FP patients prefer quick placement in a (temporary) nursing home, while PP patients opt to wait for availability at their preferred nursing home. We denote the FP group as elements of the set $P^{FP} \subseteq P$ and the PP group as elements and of the set $P^{PP} \subseteq P$.

5.5.1 Instance specification

The small setting contains four nursing homes with 20 beds each. First, we elucidate the patients' utilities and then the dynamic parameters.

For all patients $p \in P$ and nursing homes $n \in M_p \setminus 0$, we set the fixed utility to $g_{pn} = 30$, and for $n \in L_p$, $g_{pn} = M$, where M denotes a large number, e.g. 1,000. The FP and PP utility groups can be identified by their unique waiting-time utilities. For the FP group, the waiting time utilities equal $v^U_{pn}(w_p) = 0.1w_p + 100$ for $p \in P^{FP}$. This implies that in combination with $g_{pn} > 0$ for $p \in P, n \in N \setminus 0$, if a bed is available in a nursing home, a patient of the FP group will always be placed there. For the FP group, this results in a fast placement in a nursing home. For the PP group, the waiting time utilities are $v^U_{pn}(w_p) = 0.1w_p - 500$ for $p \in P^{PP}$. In this case, patients have to wait extremely long before the utility for temporary nursing homes becomes positive, which means that patients only receive a placement in their preferred nursing home. The probabilities of arrival of a patient from the FP and PP groups are indicated by p^{FP} and p^{PP}, respectively. Thus, the arrival rates for the groups are λp^{FP} and λp^{PP}.

The waiting time utilities for the preferred nursing homes are set to differ between the two groups. Namely, $v^F_{pn}(w_p) = 0.1w_p$ for $p \in P^{FP}$, and $v^F_{pn}(w_p) = 0.15w_p$ for $p \in P^{PP}$. This implies that the utilities for the preferred nursing homes of PP increase slightly faster than those of FP. In this way, at some moments, a PP group patient might surpass the waiting FP patients, although the waiting time is shorter. The utilities are chosen in this way to place PP patients faster in their preferred nursing home. Finally, the replacement penalty K is set to 1,000, which prohibits replacements to be performed between temporary nursing homes.

Dynamic parameters are set with realistic values, where available. The mean length-of-stay in nursing homes is approximately three years, reflecting real-world data [180]. Estimating the abandonment rate at home, θ, is challenging due to data scarcity. The abandonment rate at home is assumed to be higher than the service rate in nursing homes, given care limitations and the potential for patients to seek alternatives outside the region. Thus,

θ^{-1} is set to two years. In the absence of abandonments, we set $\theta = 0$. This adjustment prompts a modification in the waiting-time utility function, specifically for patients who prioritize placement in preferred nursing homes. Here, $v_{pn}^F(w_p) = 0.25w_p \; \forall \; p \in P^{PP}$, accommodating the altered dynamics in waiting times characteristic of abandonment-free scenarios.

We define the offered load per server as $\rho = \frac{\lambda}{c\mu}$. In case there are no abandonments ($\theta = 0$), we need at least that $\rho < 1$ to attain stability of the system. However, the assumption only applies to parts of our small-scale setting. For the allocation model, we allow $\rho \geq 1$ as long as $\theta > 0$. In our experiments, when studying the impact of the offered load, we will only vary the arrival rate λ and keep the service rate μ fixed. An overview of the parameter settings is given in Table 5.3.

Parameter	Value	Parameter	Value
$\lvert N \rvert$	4	K	1,000
$\lvert L_p \rvert$	$1 \; \forall \; p \in P$	$v_{pn}^U(w_p)$	$0.1w_p + 100 \; \forall \; p \in P^{FP}$
θ^{-1}	$\{\infty, 730\}$	$v_{pn}^U(w_p)$	$0.1w_p - 500 \; \forall \; p \in P^{PP}$
μ^{-1}	1,095	$v_{pn}^F(w_p)$	$0.1w_p \; \forall \; p \in P^{FP}$
M	1,000	$v_{pn}^F(w_p)$	$\{0.25w_p, 0.15w_p\} \; \forall \; p \in P^{PP}$
c_n	Initial value: $20 \; \forall \; n \in N$	p^{FP}	0.5
Δ	1 day	p^{PP}	0.5
g_{pn}	30 $\; \forall \; p \in P, n \in N$		

Table 5.3: Parameter values for the small setting.

5.5.2 Results for the small setting

First, we investigate the situation without abandonments, i.e., $\theta = 0$ and $v_{pn}^F(w_p) = 0.25w_p \; \forall \; p \in P^{PP}$. The performance of the allocation model can be compared with respect to the two extreme policies *separate* and *shared* queues (see Section 5.4.4), which is shown in Figure 5.3. It can be seen that the mean waiting time of the allocation model is much shorter than the mean waiting time of the separate queues, whereas it is only slightly longer compared to the shared queue (i.e., the lower bound). This observation is in line with the principle that "a little flexibility goes a long way"; see, e.g. [187] for SBR examples in which some flexibility in resource pooling suffices to obtain most of the efficiency gain.

To get more insight into the performance measures of the two groups of patients, we split the waiting times of the allocation model into groups, as shown in Figure 5.4. We see that for all different loads, the mean waiting time until the placement of the FP group is the lowest. In addition, the time until a patient of this group is placed in their preferred nursing home is longer than for the PP group. In that respect, the needs of both groups are indeed served: FP is placed quickly, whereas PP is placed slightly faster in their preferred nursing

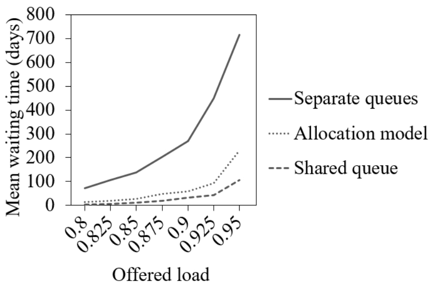


Figure 5.3: Comparison of allocation model to extremes ($\theta = 0$).

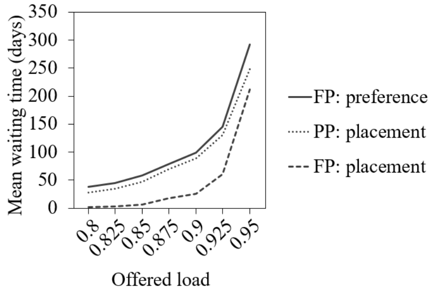


Figure 5.4: Comparison of patient groups within allocation model ($\theta = 0$).

home.

The next experiment is conducted on the effect of the replacement penalty K. Note that in our parameter setting, K only influences the replacement potential of the FP group, as the PP group is not temporarily placed. In addition, for this experiment, we diversify our g_{pn}-values into $g_{pn} \in \{10, 20, 30, 40\}$, where each value is selected with probability 0.25. The results of the different replacement penalties are shown in Figure 5.5. Clearly, an increase in K-penalty is observed to lead to fewer replacements of the FP group. In addition, we see that the number of replacements between two temporary nursing homes increases as the system is more congested. This seems in line with the fact that congestion leads to longer waiting times and consequently to more potential benefits of temporary replacements.

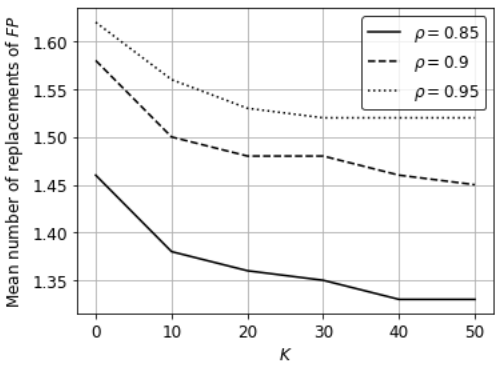


Figure 5.5: Effect of increasing K on the number of replacements of FP.

The results of the simulation model, including abandonments, are detailed in Table 5.4.

The values in Table 5.4 are the results of the simulation model that incorporates uncertainty. Since

Across various load levels, the allocation model exhibits slightly higher abandonment rates compared to the shared queue but significantly lower rates than separate queues, indicating efficient capacity utilization. The mean occupancy levels for AM are only marginally lower than those for the shared queue. The efficiency is also apparent from the mean waiting times until placement, which is only slightly higher for AM compared to the shared queue. In addition, the waiting times for the preferred AM placement are only slightly longer than those of separate queues. AM is the only model using temporary nursing home placements. However, the fraction of patients who end up in their preferred nursing home is similar to both separate and shared queue strategies. For the shared queue policy, patients are treated as having no preferences (all nursing homes are preferred), providing a clear upper bound. Naturally, the shared queue policy does not facilitate patient preferences.

Based on these results, the allocation model is found to have 'the best of both worlds': both short waiting times are obtained next to preferences being retained well (both are relatively close to their lower and upper bound, respectively). However, these achieved gains are at the cost of the number of replacements, which increases by more than 20% for all offered loads. Hence, the allocation model yields promising results only if more patient replacements are allowed.

5.5.3 Comparison to optimal solution

To compare the performance of the allocation model with the optimal solution, we developed an MDP that gives us the optimal policy for small instances. More information on the MDP setup can be found in Appendix 5.A.2. We ran an experiment with 50 randomly generated small instances, where we have two nursing homes with two beds each. We chose the parameters as much as possible in line with previous experiments. For example, we choose μ such that the maximum total departure rate is equal to that of the small-scale example, i.e., we let $\mu = \frac{80}{4} \cdot \frac{1}{1095} = 0.018$, where the first ratio is the number of beds in the small-scale example divided by the number of beds in the current experiment. To create diversity in the numerical experiments, we added randomization and let $\mu \sim$ Unif$(0.01, 0.03)$. All parameter values are realizations of the values provided in Table 5.5. Note that the notation in the table is intuitively clear, but the formal definitions can be found in Appendix 5.A.2.

To align the allocation model with the optimal solution, we use the same classes, g_{in} values, and similar waiting utility, that is, $v_{in}^U(w) = v_{in}^F(w) = w_i w$, $\forall\ i \in I, n \in N$ for the generated w_i in the instances. This ensures that the waiting costs for the MDP correspond to the increase in waiting utility in the allocation model. The resulting long-term average rewards for the allocation model are determined through simulation, using the parameter values described in Table 5.5, to ensure consistency with the optimal solution found by the MDP. The optimality gap between the long-term average reward of the allocation model

the model ran for a sufficient duration, the uncertainty around the results is negligible and therefore omitted.

Policy	% abandonments	% died at temp. NH	% died at pref. NH	MWT till placement (d)	MWT till preferred (d)	Mean nr. of replacements	Mean occupancy
Offered load: 0.9							
Shared	1.5%	0.0%	98.5%	10.3	10.3	0.98	0.89
AM: tot	2.8%	2.5%	94.6%	18.4	44.2	1.14	0.87
AM: FP	0.6%	5.1%	94.3%	4.3	55.5	1.35	-
AM: PP	5.1%	0.0%	94.9%	33.1	33.1	0.95	-
Separate	5.8%	0.0%	94.2%	38.4	38.4	0.94	0.85
Offered load: 1							
Shared	5.2%	0.0%	94.8%	37.7	37.7	0.95	0.95
AM: tot	6.3%	3.5%	90.2%	41.2	78.3	1.18	0.94
AM: FP	2.5%	7.0%	90.5%	17.4	89.6	1.47	-
AM: PP	10.0%	0.0%	90.0%	67.0	67.0	0.90	-
Separate	9.7%	0.0%	90.3%	65.6	65.6	0.90	0.90
Offered load: 1.1							
Shared	10.2%	0.0%	89.8%	75.8	75.8	0.90	0.98
AM: tot	11.5%	3.3%	85.1%	76.3	115.4	1.14	0.97
AM: FP	6.7%	6.7%	86.6%	46.4	121.0	1.44	-
AM: PP	16.3%	0.0%	83.7%	109.7	109.7	0.84	-
Separate	14.6%	0.0%	85.4%	101.6	101.6	0.85	0.94

Table 5.4: Results for the small setting.

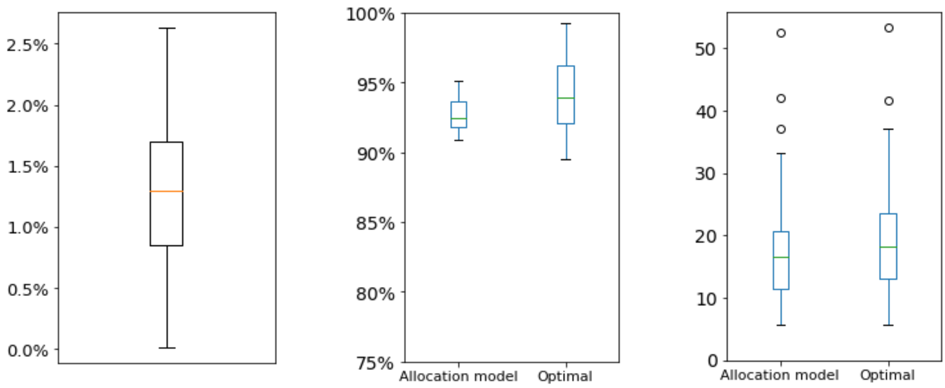
Parameter		Value from distribution
μ		Unif$(0.01, 0.03)$
$\lambda = p_\lambda \mu \sum_{n \in N \setminus \{0\}} b_n$		with $p_\lambda \sim$ Unif$(0.5, 0.99)$
g_{in}	$i \in \{FP_1, FP_2\}, n \in M_i$	Unif$(20, 40)$
g_{in}	$n \in L_i$	Unif$(50, 70)$
K_i	$i \in \{FP_1, FP_2\}, n \in N \setminus \{0\}$	Unif$(1, 10)$
w_i	$i \in \{FP_1, FP_2\}$	Unif$(0.5, 1.5)$
w_i	$i \in \{PP_1, PP_2\}$	Unif$(1.5, 3)$

Table 5.5: Parameter values for randomly generated instances.

(denoted g^{AM}) and the optimal solution (denoted g^*) can now be defined as

$$\frac{g^{AM} - g^*}{g^*}.$$

For the 50 instances, we observe a mean optimality gap of 1.6%, which is rather small. Figure 5.6a illustrates the distribution of this gap. Furthermore, Figure 5.6b shows the percentages of patients who died in preferred nursing homes, while Figure 5.6c presents the waiting time distributions. Although the allocation model slightly lags behind in "accurate placement" (with an average of 92.7% compared to 94.2% for the MDP solution), it compensates with shorter waiting times. The mean waiting time for the allocation model is marginally lower at 17.6 days, attributed to its immediate placement policy once a suitable bed is available.



(a) Optimality gap in %

(b) % died in pref. NH

(c) Waiting time (d)

Figure 5.6: Results of the small instances.

We conducted 50 instances of a slightly larger MDP, maintaining the parameter values from Table 5.5 but with three beds per nursing home instead of two. The mean optimality gap for this setting was 1.0%, suggesting a potential decrease in the gap as the problem scale increases (from 1.6% to 1.0%). These additional results are detailed in Appendix 5.A.3.

Overall, the allocation model shows close-to-optimal performance with a mean optimality gap of 1.3%. Several factors contribute to its excellence. First, the ability to relocate patients post-placement allows for rectification of undesirable placements, which benefits static or greedy policies (cf. [94]). Second, the structure of waiting-time utilities exhibits similarities to threshold policies, known for their optimality in dynamic systems [19, 100]. Third, uniformity in length of stay and abandonment rate between patient types creates a more homogeneous system. Lastly, it is worth noting that static policies have been proven optimal in various multi-class queue scenarios, such as the well-known $c\mu$ or $c\mu/\theta$ rule [13].

5.5.4 Effect of multiple preferred nursing homes

In the small setting, the number of nursing homes preferred by all patients is set to one. However, patients may be amenable to selecting multiple nursing homes as their preferred option. For that reason, we show the effect of choosing multiple nursing homes as their preferred ones. The results for the small setting without abandonments are presented in Figure 5.7.

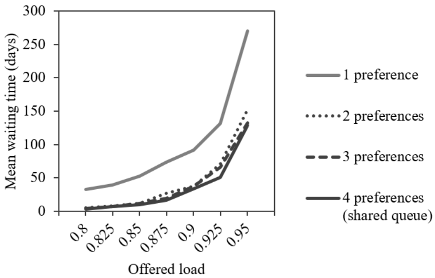


Figure 5.7: Mean waiting time AM multiple preferences ($\theta = 0$).

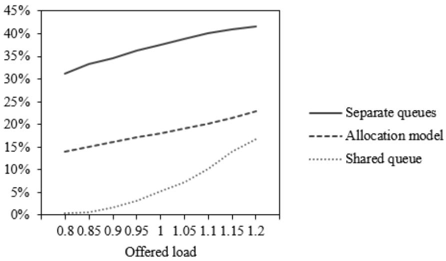


Figure 5.8: Percentage abandonments scenario popular nursing homes.

It becomes clear that if patients select two preference homes, the waiting times for various offered loads are nearly as short as the (minimal) waiting times of the shared queue. In this setting, increasing the number of preferred nursing homes to two is a very effective intervention, which is in line with the principle that adding a little flexibility provides most of the benefits of complete flexibility [91, 187].

5.5.5 Effect of popular nursing homes

Although the preceding results are obtained for nursing homes with equal demand, in practice some nursing homes are more popular than others. Therefore, we define a scenario in which demand is not equally distributed. In this scenario, a nursing home $j^* \in N$ is selected with probability 0.5 as the preferred nursing home for each patient $p \in P$, hence $P(l_p := j^*) = 0.5$. The comparison between the number of abandonments under the allocation model and the extremes is shown in Figure 5.8.

First, note that the policy for the shared queue does not change, since in this policy we do not take into account individual preferences. In contrast, we see in Figure 5.8 that the percentage of abandonments for the separate queue policy and the allocation model has increased significantly. Compared to the results without popular nursing homes, the abandonment fraction $\rho = 1$ has increased for the separate queues from 9.7% to 37.7% and for the allocation model from 6.3% to 18.0%. We thus see that the allocation model is far better in retaining the waiting times, and therefore fraction of abandonments, at an acceptable level.

5.6 Conclusion and Discussion

The allocation model proposed in this chapter offers an alternative to conventional waiting lines for nursing home placements, boasting a patient-centered approach. By prioritizing individual preferences and allowing flexibility in waiting times, the model achieves a balance between minimizing waiting times, similar to shared queues, and enabling patients to choose their preferred nursing home, as if separate queues were employed. Moreover, the optimal placement policy found by solving an MDP shows that, for small instances, the mean optimality gap of the allocation model equals 1.3%. Hence, the model is found to have 'the best of both worlds', resulting in a quicker placement of patients into the nursing home of their preference.

Future research avenues include incorporating predictive information about bed availability to enhance allocation accuracy and expanding the model's scope to encompass broader elderly care contexts, such as addressing bed-blocking in hospitals or exploring the impact of transitional care units. Furthermore, the promising outcomes of the allocation model underscore the importance of centralized monitoring of nursing home waiting lists. In privatized care settings such as the Netherlands, where nursing homes operate independently, information sharing and collaborative strategies in the elderly care sector can significantly enhance efficiency for both patients and nursing homes. A potential direction towards practical implementation is to investigate the set-up of a regional care center that manages the available beds, waiting list, and placement of patients, using a scientifically-based allocation model. In addition, refinement of the utility functions warrants further investigation to ensure that they align with the diverse needs of patients. These steps seem to be important prerequisites for a successful practical implementation.

Although our research focuses primarily on the elderly care domain, the allocation policy framework is relevant for various other domains facing bed scarcity and trade-offs between preferences and waiting times. From long-term care for mentally disabled people to psychiatric patient care, applying logistic perspectives to address excessive waiting times is imperative.

5.A Appendix

In this appendix, more details can be found about the simulation model, Markov Decision Process, and some additional results.

5.A.1 Simulation model

In this section, we discuss the simulation procedure of the allocation heuristic in more detail. First, denote I^t as the set of patients arrived and not yet left before time $t \in T$, where we see that the p^{th} arrival is denoted as patient $p \in I^t$. We define the patient sets that reside at time $t \in T$ at a certain location:

$$F_n^t = \{p \mid p \in I^t, l_p \in L_p\} \qquad \forall t \in T, n \in N, \tag{5.6}$$
$$G_n^t = \{p \mid p \in I^t, l_p \in M_p\} \qquad \forall t \in T, n \in N. \tag{5.7}$$

Thus, at time $t \in T$, F_n^t denotes patients who reside permanently in (final) nursing home $n \in N$. G_n^t is the set with temporary patients at nursing home $n \in N$. Note that G_0^t is the set of patients at home.

The simulation procedure can be described as follows; see Figure 5.9 for a schematic representation. First, the initialization takes place, which includes a warming-up period for the model. The warming period begins with adding patients to nursing homes until an occupancy of 90%; this is to speed up the process of reaching steady state. To determine if and when the steady state was reached, we used a graph with confidence intervals of the waiting times until placement. For the small instances, the warming period was ended after 1,000 clients had left the system, and for the full-size instances that were used for the results in Chapter 6, this was reached after 10,000 clients had left.

Then, in the dynamic setting, the time is set to the next time moment (5.8) and the system is updated according to all events that occurred during the interval of length Δ. The corresponding output values are updated. Then, for the current $t \in T$ the sets I^t, G_n^t and F_n^t $\forall\, n \in N$ are defined in (5.14). The set of patients P is updated in a way that includes new arrivals and excludes departures (5.15). The capacity of all nursing homes is updated so that it includes only empty beds and beds occupied by temporary patients (5.16). At the end of the dynamic setting, the utilities are updated according to the new waiting times (5.17).

In the next step, the allocation model is run (5.18). Finally, the resulting replacements are carried out, including an update of the locations l_p of all patients $p \in P$ (5.19).

Finally, two stopping criteria were defined on the basis of the instance size. For the toy examples, the simulations were run until the width of the confidence interval for waiting times was less than 1% of the mean waiting time for placement. For realistic full-size instances, this was not achievable, and therefore we ran the simulations for at least twelve hours and presented the results accompanied by their confidence intervals.

5.A.2 MDP for optimal allocation

Let X_t denote the state of the system at times $t \geq 0$, defined in the following, such that $\{X_t\}_{t \in T}$ represents the stochastic process. Moreover, let Π denote the set of allocation policies. Then, policy $\pi \in \Pi$ has a long-term average reward

$$r_\pi = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} r_\pi(X_t).$$

The optimal policy is a policy π for which we have $max_\pi r_\pi$.

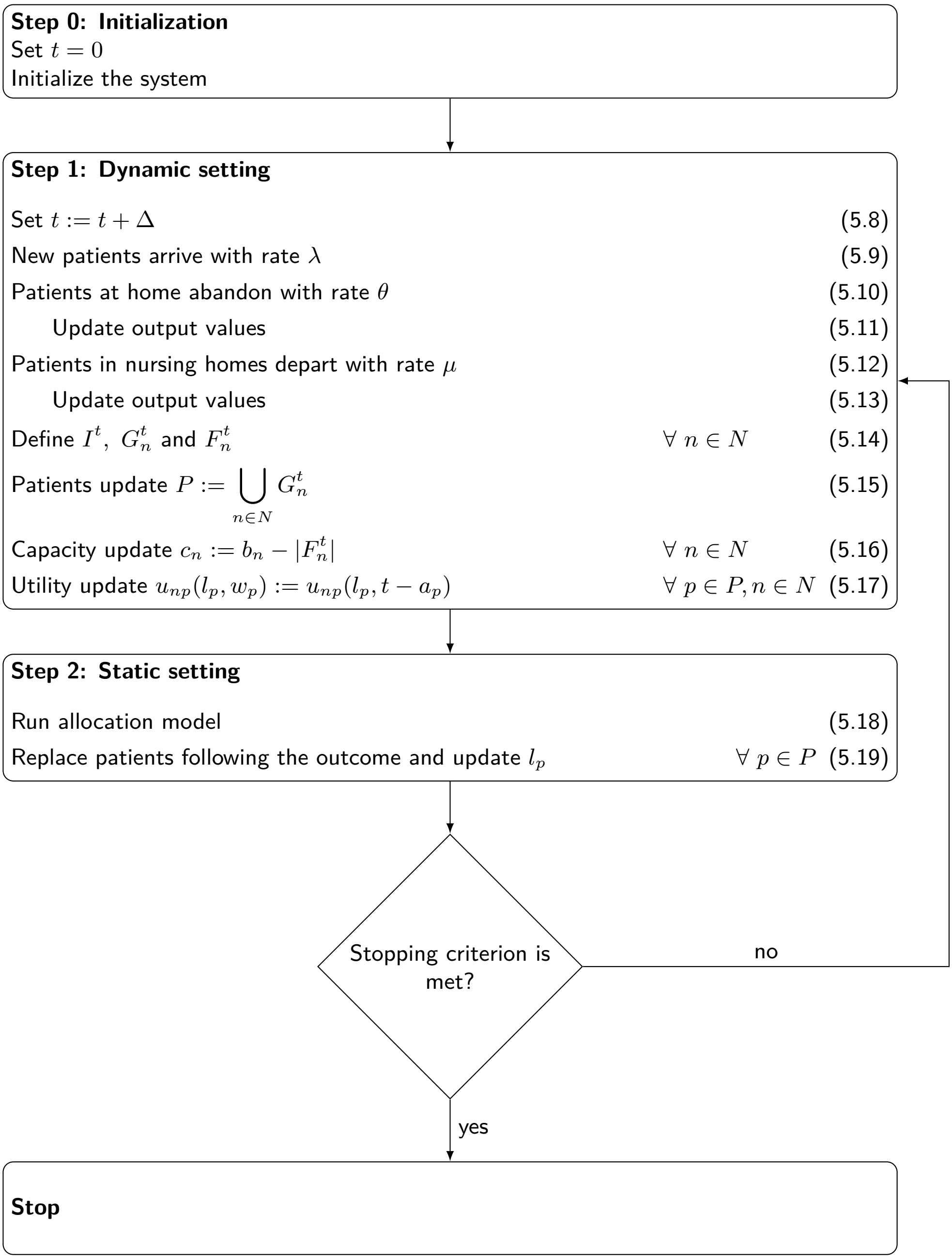


Figure 5.9: Overview of the simulation procedure.

Problem formulation

Now we discuss all components of the MDP: the state space, action space, transition probabilities, and rewards. All notation used in this section can be found in Table 5.6.

Sets	
I	Patient classes
N	Nursing homes, with $n = 0$ the home location
$L_i \subseteq N$	Preferred nursing homes of class $i \in I$
$M_i \subseteq N$	Non-preferred nursing homes of class $i \in I$, i.e. $N \setminus L_i$
Parameters	
λ_i	Arrival rate of class $i \in I$
μ	Service rate
r_{inm}	Lumpsum reward of class $i \in I$ residing at location n moving to location m
X_{in}	The number of (temporary) patients of class $i \in I$ at location $n \in N$
y_n	The number of preferred persons at location $n \in N$
w_i	Waiting penalty of class $i \in I$
c_{rej}	Rejection costs
b_n	Number of beds available at location $n \in N$
A_{imn}	Number of patients of class $i \in I$ allocated from location m to location $n \in N$
g_{in}	(Initial) willingness of class $i \in I$ to be placed in nursing home $n \in N$
K_i	Replacement penalty of class $i \in I$

Table 5.6: MDP notation.

State space

In the allocation model, we have that each patient is unique. In an MDP framework, this leads to an excessive state space for realistic model instances. Therefore, we define a set of patient classes I for the MDP. The patients of class $i \in I$ arrive according to a Poisson process with rate λ_i, have sets of preferred nursing homes L_i and non-preferred nursing homes M_i.

For the state space, we introduce the variable $X_{in} \in \mathbb{N}$ defined as the number of patients of class $i \in I$ residing in nursing home $n \in N$. Let $\mathbf{X} \in \mathbb{N}^{|I| \times |N|}$ be the matrix with the entries X_{in}. In addition, we only need to keep track of the classes of patients that are not yet in their preferred nursing home, since only those are candidates to be replaced. For the preferred placed patients (that is, patient classes i for which $n \in L_i$), only the number of patients $y_n \in \mathbb{N}$ satisfies. We store those elements in a vector $\mathbf{y} \in \mathbb{N}^{|N|}$, with y_n as the nth entry. Note that we truncate the total number of patients on the waiting list at b_0, such that the state space remains finite. This implies that patients arriving who find b_0 patients on the waiting list are rejected.

Hence, a state can be described as $\{\mathbf{X}, \mathbf{y}\}$, where the state space is given by

$$S = \left\{ \{\mathbf{X}, \mathbf{y}\} \mid \sum_{i \in I} X_{in} + y_n \leq b_n \ \forall n \in N \right\}.$$

Action space

As actions we define the number of temporarily placed patients of a certain class $i \in I$ that are replaced from nursing home m to $n \in N$, denoted by A_{imn}. We define the action matrix $\mathbf{A} \in \mathbb{N}^{|I| \times |N| \times |N|}$ with elements A_{imn}. Note that we allocate all temporary patients, which can also be in their current location. Now, let $Y(\{\mathbf{X}, \mathbf{y}\})$ denote the set of all possible actions for state $\{\mathbf{X}, \mathbf{y}\}$. Then, we have

$$Y(\{\mathbf{X}, \mathbf{y}\}) = \left\{ \mathbf{A} \mid X_{in} = \sum_{m \in N} A_{inm}, \ \sum_{i \in I} \sum_{m \in N} A_{imn} + y_n \leq b_n \ \forall n \in N \right\}.$$

Transition probabilities

The transition probabilities consist of two 'consecutive' parts. First, based on the actions to allocate patients to other locations, we move from state $s \rightarrow s^*$. Then, based on probabilities induced by the stochastic process, we arrive at our final state s'.

For the first part, we have the following state transitions from $\{\mathbf{X}, \mathbf{y}\}$ to $\{\mathbf{X}^*, \mathbf{y}^*\}$:

$$\begin{aligned} X^*_{in} &= \sum_{m \in N} A_{imn}, && \text{for } n \in M_i, \\ y^*_n &= y_n + \sum_{m \in N} A_{imn}, && \text{for } n \in L_i. \end{aligned}$$

Now consider the second part, that is, the transition probabilities out of the state s^*. Without loss of generality, we rescale the time such that $\sum_{i \in I} \lambda_i + \sum_{n \in N \setminus \{0\}} b_n \mu = 1$, in which case the rates can be interpreted as transition probabilities. Moreover, for notation purposes, we define the single-entry matrix E_{ij} as a matrix with zeroes and a 1-entry in the ith row and the jth column, and the unit vector e_j as a vector with zeroes and a 1 as the jth element. The final transition probabilities are

$$P_{(\{\mathbf{X},\mathbf{y}\}),(\{\mathbf{X}',\mathbf{y}'\})}(\mathbf{A}) =$$
$$\begin{cases} \sum_{i\in I}\lambda_i & \text{if } \mathbf{X}'=\mathbf{X}^*,\ \mathbf{y}'=\mathbf{y}^*, \\ & \sum_{i\in I} X^*_{i0}=b_0 \\ \lambda_j & \text{if } \mathbf{X}'=\mathbf{X}^*+E_{j0},\ \mathbf{y}'=\mathbf{y}^*, \\ & \sum_{i\in I} X^*_{i0}<b_0,\ j\in I \\ \mu X^*_{jm} & \text{if } \mathbf{X}'=\mathbf{X}^*-E_{jm},\ \mathbf{y}'=\mathbf{y}^*, \\ & j\in I,\ m\in N\setminus\{0\} \\ \mu y^*_m & \text{if } \mathbf{X}'=\mathbf{X}^*,\ \mathbf{y}'=\mathbf{y}^*-e_m, \\ & m\in N\setminus\{0\} \\ 1-\mu\left(\sum_{i\in I}\sum_{n\in N\setminus\{0\}} X^*_{in}+y^*_n\right)-\sum_{i\in I}\lambda_i & \text{if } \mathbf{X}'=\mathbf{X}^*,\ \mathbf{y}'=\mathbf{y}^*. \end{cases}$$

Rewards

The reward at each decision epoch depends on the action reward r_{imn} (defined below), waiting costs w_i and rejection costs c_{rej}. The formula for this is the following:

$$r(\{\mathbf{X},\mathbf{y}\},\mathbf{A}) = \sum_{i\in I}\sum_{m\in N}\sum_{n\in N} r_{imn}A_{imn} - \sum_{i\in I} w_i X_{i0} - c_{rej}\sum_{i\in I}\lambda_i \mathbb{1}_{\{\sum_{i\in I} X^*_{i0}=b_0\}}.$$

The action rewards are based on the utilities for placement in nursing homes, as provided in (5.1), without the waiting times utilities. For notational convenience, we set $g_{pn} := g_{in}$ for the patients p that belong to class i. The precise relation between the utilities and the action rewards r_{imn} is given by

$$r_{imn} = \begin{cases} g_{in} & \text{if } m=0,\ n\in N\setminus\{0\}, \\ g_{in}-g_{im}-K_i & \text{if } m\in M_i,\ n\in L_i, \\ 0 & \text{if } m=n, \\ -\infty & \text{otherwise.} \end{cases} \tag{5.20}$$

Value iteration

In order to solve the MDP and find the optimal long-term average reward g^*, we use value iteration. Hence, we need to find the value function $V(s)$ for all states $s \in S$. This is done by iteratively computing the value function $V_n(s)$ until convergence occurs. For this purpose, we initialize $V_0(s) = 0\ \forall s \in S$ and then iteratively solve the Bellman equations for $i \in S$,

$$V_{n+1}(i) = \max_{a \in A(i)} r(i,a) + \sum_{j \in S} P_{ij}(a) V_n(j).$$

After each iteration, we compute $h_n = \inf_i |V_n(i) - V_{n-1}(i)|$ and $H_n = \sup_i |V_n(i) - V_{n-1}(i)|$. The value functions are converged if $H_n - h_n < \epsilon h_n$, where ϵ is a prespecified accuracy. As the value function has converged, we have $\forall\ s \in S$, $V(s) = V_n(s)$ and $g^* = V_n(s) - V_{n-1}(s)$.

State space size

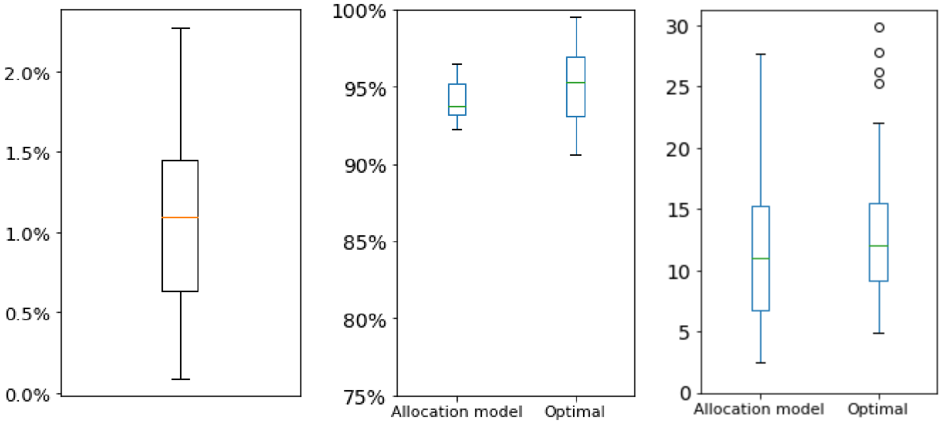
To illustrate the impact on the state space, we define the following instances to compare the allocation model with the optimal solution. We take two nursing homes, $N = \{0, 1, 2\}$ with 0 the home location, and four classes, $I = \{\mathsf{FP}_1, \mathsf{FP}_2,$
$\mathsf{PP}_1, \mathsf{PP}_2\}$. The class subscripts correspond to the preferred nursing home. For this example, the number of possible states for $\mathbf{X}$ and $\mathbf{y}$ provides an upper bound for the size of the state space:

$$|\mathbf{X}| = \binom{b_0 + 4}{b_0} \binom{b_1 + 2}{b_1} \binom{b_2 + 2}{b_2},$$
$$\text{and } |\mathbf{y}| = (b_1 + 1)(b_2 + 1).$$

As can be seen, the size of the state space grows quickly. Moreover, we see that the number of possible actions is also large, since the empty beds can be filled by all possible combinations of waiting patients. For these reasons, we are only able to solve small instances.

5.A.3 Additional results MDP instances

In this appendix, we present the results for the instances defined in Section 5.5.3, with 3 beds per nursing home. As can be seen in Figure 5.10a, the optimality gap is lower for these instances than for the instances with 2 beds per nursing home. For the other output measures, the results for the instances with 3 beds show a similar behavior to the instances with 2 beds per nursing home; see Figures 5.10b and 5.10c.



(a) Optimality gap in % **(b)** % died in pref. nursing home **(c)** Waiting time (d)

Figure 5.10: Results additonal instances

6 Allocation of Older Persons to Nursing Homes: Practice

Based on:

Arntzen, R. J., Bekker, R., Smeekes, O. S., Buurman, B. M., Willems, H. C., Bhulai, S., & van der Mei, R. D. (2022). Reduced waiting times by preference-based allocation of patients to nursing homes. *Journal of the American Medical Directors Association*, *23*(12), 2010–2014.

Arntzen, R. J., Bekker, R., & van der Mei, R. D. (2024b). Preference-based allocation of patients to nursing homes. *Operations Research for Health Care*, *42*, 100442.

Abstract

Objective: The long waiting times for nursing homes can be reduced by applying advanced waiting-line management. In this chapter, we discuss the preference-based allocation model for older adults to nursing homes, evaluate the performance in a simulation setting for three case studies, from two urban regions (Amsterdam and Rotterdam) and one rural region (Twente), and elaborate on the implementation in practice.

Method: Data about somatic nursing homes and capacities for the three case studies were identified. A set of preference profiles was defined with aims regarding waiting time preferences and flexibility. Guidelines for implementation of the model in practice were obtained by addressing the tasks of all stakeholders. Thereafter, the simulation was run to compare the current practice with the allocation model based on specified outcome measures about waiting times and preferences. First, for Amsterdam, the current policy was validated using waiting times data, and a sensitivity analysis was run. Second, for Rotterdam and Twente, more elaborate preference profiles were tested.

Results: We found that the allocation model decreased the waiting times for the three case studies. Compared with the current practice policy, the allocation model reduced the waiting times until placement by at least a factor of two (from 313 to 85 days in Amsterdam, from 166 to 80 days in Rotterdam, and from 178 to 82 days in Twente). Moreover, more of the older adults ended up in their preferred nursing home and the aims of the distinct preference profiles were satisfied.

Conclusion: The results show that the allocation model outperforms commonly used waiting-line policies for nursing homes, while meeting individual preferences to a larger extent. Moreover, the model is easy to implement and of a generic nature and can, therefore, be extended to other settings as well (e.g., to allocate older adults to home care or daycare). Finally, this research shows the potential of mathematical models in the care domain for older adults to face the increasing need for cost-effective solutions.

6.1 Introduction

In the previous chapter, an allocation model for patients to nursing homes was introduced. It was found that by applying the allocation model, the older adults' waiting times were significantly reduced, whereas at the same time their personal preferences were served better. In this chapter, we aim to extend the research by addressing the allocation model from a practical perspective. The allocation model is preference-centered and therefore a crucial aspect is to make sure the individual preferences are correctly obtained. We evaluate the allocation model using newly developed preference profiles on three real-life case studies. Hence, the following research questions are answered in this chapter:

1. How can the allocation model be implemented in practice?
2. How does the allocation model perform for the different preference profiles?

6.2 Method

We first sketch the background of the study, then describe the allocation model and discuss aspects of the implementation in practice. For more information about the (mathematical) set-up of the model, we refer to the previous chapter.

The preference profiles were defined in the following way. We first specified the flexibility of the older adults: more specifically, we defined that older adults may select one or two preferred nursing homes. The more preferred nursing homes, the more flexibility the older adult offers, which leads to a shorter (average) waiting time. Moreover, for the waiting time preference, recall that we formulated two options as well: the type of older adults who want a fast placement and the type of older adults who want a preferred placement. The type Fast Placement (FP) wants to be placed as fast as possible in a (temporary) nursing home, since the older adult's situation at home is unlivable (e.g., acute care patients). On the other hand, the type Preferred Placement (PP) only wants to be placed in a nursing home that is one of the preferred nursing homes, and otherwise prefers to stay at home. After selecting both the flexibility number of preferred nursing homes and the waiting time preferences, the four preference profiles were obtained, as provided in Table 6.1.

In order to evaluate the allocation model and the current practice model in terms of the outcome measures, a simulation study was performed. This was done for the three regions Amsterdam, Rotterdam, and Twente. Results are given in total days, calculated in means accompanied by confidence intervals. The 95%-confidence intervals indicate the probabilistic bounds on the statistics, i.e., if two values show non-overlapping confidence intervals, the statistics are significantly different [98].

Abbreviation	Preference Profile	Explanation
FP1	Fast Placement 1	Older adult wants to be placed fast in a nursing home and has one preferred nursing home
PP1	Preferred Placement 1	Older adult wants to be placed in a preferred nursing home and has one preferred nursing home
FP2	Fast Placement 2	Older adult wants to be placed fast in a nursing home and has two preferred nursing homes
PP2	Preferred Placement 2	Older adult wants to be placed in a preferred nursing home and has two preferred nursing homes

Table 6.1: Abbreviations and preference profiles.

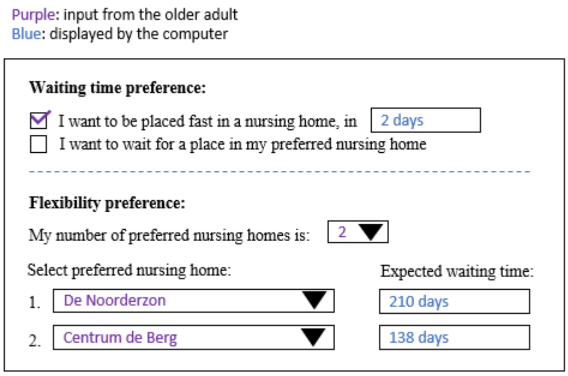
6.2.1 Implementation in practice

Besides evaluating the allocation model by simulation, we also formulate guidelines for the implementation of the allocation model in practice. For this purpose, we identified the three stakeholders that are involved, namely a *placement office*, *older adults*, and the nursing homes' *logistic managers*. We describe here the needs and requirements of all stakeholders in order to sketch the use of the allocation model in practice.

First, a placement office needs to be set up or an existing institution must be appointed to take on this role. This office should organize the placement of older adults in nursing homes for the surrounding region. The placement office must have a digital infrastructure that stores older adults' applications and nursing home capacities, and uses this as input for the allocation model. The placement office runs the allocation model at structural time moments, for example, every morning. Then, the placement office needs to inform the older adults and the logistic managers about the relocations that are the outcomes of the allocation model.

The next stakeholders are the logistic managers that are currently involved with monitoring the waiting lists of their nursing homes. After implementation of the allocation model, this task is transferred to the placement office. The new task of the logistic managers is to provide the placement office with real-time data about the capacities (i.e., free beds and temporarily placed older adults), such that the allocation input remains up-to-date.

Finally, the older adults need to express their preferences for nursing homes to the placement office. For that purpose, the older adults need to fill in a, preferably digital, preferences form, as provided in Figure 6.1. As one can see, an older adult can choose whether he/she wants to be placed as fast as possible in a temporary nursing home, which is in expected two days in the example, or wants to wait at home. Moreover, the older adult can select how many - and which - nursing homes are chosen as preferred ones, and subsequently the expected waiting times are displayed automatically. When these projected waiting times are too long, the older adult may reconsider his/her preferences (i.e., be more flexible) to obtain lower waiting times. This way, the older adult can interact with the system until the older adult is satisfied with the outcome.


Purple: input from the older adult
Blue: displayed by the computer

Waiting time preference:

☑ I want to be placed fast in a nursing home, in 2 days
☐ I want to wait for a place in my preferred nursing home

Flexibility preference:

My number of preferred nursing homes is: 2

Select preferred nursing home:	Expected waiting time:
1. De Noorderzon	210 days
2. Centrum de Berg	138 days

Figure 6.1: Example of the digital preferences form.

6.2.2 Current policy

The current allocation policy in the Netherlands allows patients to sign up for one preferred nursing home and wait at home until placement [86]. However, if the waiting time becomes excessive, patients may opt for immediate placement due to impatience. As bed availability is not centrally managed and information is limited [86], a regional office manager contacts nursing homes individually to inquire about temporary residency options, ceasing calls in order to find a suitable bed. Under this policy, patients are initially placed on a waiting list for their preferred nursing home and later added to a secondary waiting list shared among all regional nursing homes, where beds are allocated if the primary list is empty. This policy framework is depicted in Figure 6.2.

We stress that the policy introduced here is the current *policy*, which may differ from the current *practice*. Namely, no centralized administration system exists that prohibits patients to register for multiple nursing homes, although this is against the rules. This phenomenon is recognized by experts as the 'grey waiting list'.

6.3 Case study: Amsterdam

To show the results of the model in a realistic parameter setting, we applied the model to the situation in Amsterdam, the capital of the Netherlands. First, we describe the obtained data, after which we validate our model by validating the assumptions and using information on current waiting times. Thereafter, the results of our proposed alternative - the allocation model - are discussed.

Note that all patients initially enter one queue. Then, after waiting $>$ 15 months, the patient additionally takes place in the common queue with second priority, which is displayed for patient A and B.

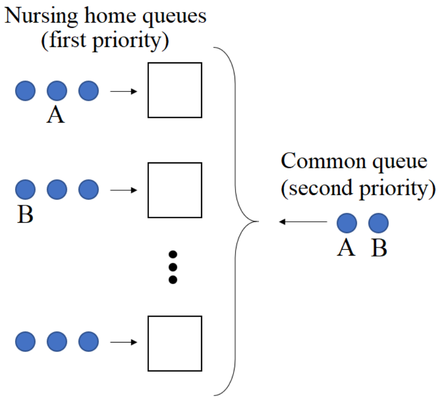


Figure 6.2: Illustration of the current policy model.

6.3.1 Data

For our case study, we focus on Amsterdam and a specific patient group: those with severe somatic conditions, categorized as ZZP 6 and ZZP 8 in the Dutch healthcare system [148]. These patients exhibit physical symptoms and require intensive care. In Amsterdam, there are 39 nursing homes with dedicated somatic departments, accommodating these patients, see Figure 6.3. Utilizing non-public microdata from Statistics Netherlands, we derive essential parameters. The arrival rate ($\lambda = 1.24$/day) is determined from nursing home allowance requests, while the capacity ($c = 775$) reflects the maximum number of simultaneous residents. Considering a balanced distribution of capacity across nursing homes, each is assumed to have 20 beds, resulting in 780 beds. The average length-of-stay ($\frac{1}{\mu} = 666$ days) is obtained from patient declaration data. With the offered load calculated at $\rho = 1.07$, we adjusted the arrival rate to 1.25 accordingly so that the offered load remained the same.

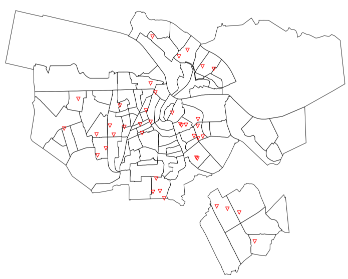


Figure 6.3: Locations of somatic nursing homes in the Amsterdam region.

For the simulations, patient arrivals are generated using demographic data from data.amsterdam.nl. The total number of people in Amsterdam older than 65 is circa 110,000. Moreover, we identify the number of people older than 65 on neighborhood level. Then, for each neighborhood we calculate the fraction f_r of the total number of elderly people who live in the specific neighborhood $r \in R$. We generate patients for neighborhood $r \in R$ with rate λf_r. The fractions of each neighborhood are shown in Figure 6.4.

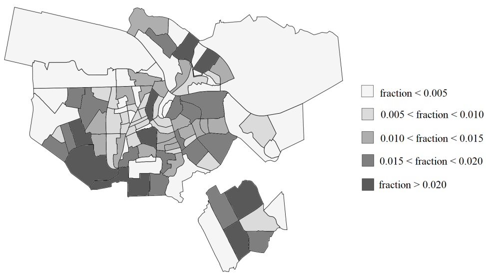


Figure 6.4: Fractions for each neighborhood in the Amsterdam region.

The preference groups correspond to the Dutch 'waiting classification' of patients: the group FP corresponds to the 'Active Placement' group and PP corresponds to 'Wait for Preference' group [2]. From this report, we know that in October 2022, 20.1% of the patients on a waiting list were of the FP category. However, as we do not know the waiting times of the different subgroups, we cannot calculate their arrival probabilities. Due to the lack of this information, we roughly set the arriving patient probabilities to $p^{PP} = 0.5$ and $p^{FP} = 0.5$.

6.3.2 Model validation

We aim to validate our model in two ways. First, we use the real-life data to verify the assumptions that we provided in the dynamic model. Thereafter, we show that our model of the current policy is a decent representation of reality, by comparing the resulting waiting times of our model by existing data on waiting times.

Assumptions validation
We validate the assumptions of our dynamic model using real-world data. Firstly, we assess the assumption that the length-of-stay in nursing homes follows an exponential distribution. Survival analysis is employed on data comprising length-of-stay, accounting for censored data where patients remain alive beyond the observation period. The resulting Kaplan-Meier curve, depicted in Figure 6.5, reveals that after three years, approximately 15% of patients are still present. A two-sample Kolmogorov-Smirnov test compares this curve with an exponential distribution fitted to the lowest 85% of values. With a p-value exceeding .01, we cannot reject the null hypothesis, indicating similarity between the dis-

tributions. Thus, we conclude that the length-of-stay for this patient group conforms to an exponential distribution.

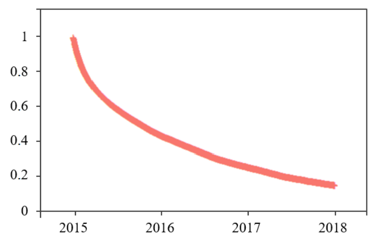


Figure 6.5: Kaplan-Meier curve: survival probability.

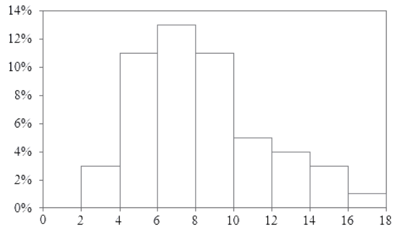


Figure 6.6: Histogram of arrival data per week in 2016.

The second assumption that we aim to validate is that the arrival process is Poisson. As data source for the arrival process, we use acceptance data for ZZP 6 and 8 indications, i.e. after acceptance patients can apply for a nursing home. A histogram with the number of arrivals per week is displayed in Figure 6.6. Moreover, a chi-square test indicated that the number of arrivals per week could be according to a Poisson distribution $\chi^2(df = 12, N = 52) = 24.8$, $p > .01$.

Current policy validation

To validate the results of the current policy model, data about waiting lines for long-term care are used. The best data source found contains information on the waiting time distribution of all patients to receive long-term care in the Netherlands [201], and thus not specified for our somatic patient population, although it is noted that "the individual indications show a similar distribution" [201, p. 1348]. The data used is from April 1, 2020 and thus pre-COVID-19. We use this information to compare to our current policy model, as displayed in Figure 6.7.

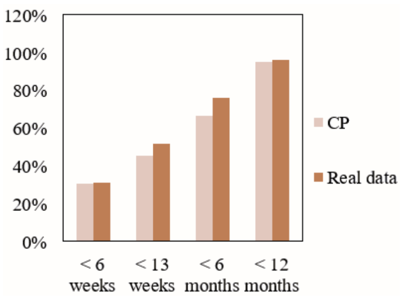


Figure 6.7: Fraction of patients waiting for some amount of time.

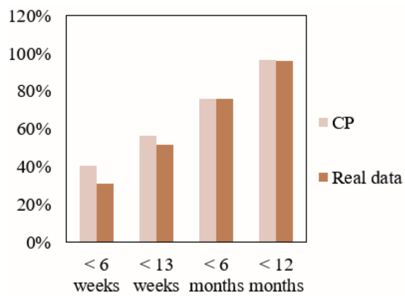


Figure 6.8: Fraction of patients waiting, with $P(3 \text{ preferred NHs}) = 0.1$.

From Figure 6.7, we conclude that the waiting times resulting from the current policy model are similar to the real-life data. However, it was found that the occupancy level resulting from the current policy model is 0.76, which is unrealistically low. Therefore, we also run a scenario that represents the current *practice* better, which is the scenario in which (against the rules) 10% of the patients register for three nursing homes instead of one. In this case, the waiting times can also be approximated fairly well, as can be seen in Figure 6.8. Next to that, the occupancy level is found to be the realistic value of 0.93. Hence, we validated the current policy by introducing a more realistic scenario for the current practice. In the following, we will use the current *policy* model as reference, since this model contains the least assumptions and is the official protocol.

6.3.3 Parameter setting for the allocation model

For the allocation model, we have defined the utilities for the patients in close collaboration with experts from the elderly care domain. To obtain the initial utilities g_{pn} for the patients to the nursing homes, we assume that preferences for certain nursing homes mainly depend on the travel distance between the home location of the patient and the nursing home, dist_{pn}, where the distance is defined as travel distance *by car*. This was motivated by research on nursing home selection where *location* was the "single most frequently cited factor" [158]. The home locations of the patients are chosen as the midpoints of the neighborhood from which they were generated. The midpoints of the neighborhoods are calculated in the following way:

$$\text{Midpoint latitude} = \frac{1}{2}(\text{maximum latitude - minimum latitude}),$$
$$\text{Midpoint longitude} = \frac{1}{2}(\text{maximum longitude - minimum longitude}).$$

Next, we used the following utility scheme for the region of Amsterdam for patient $p \in P$ towards nursing home $n \in N$:

$$g_{pn} = \begin{cases} 100 & \text{if dist}_{pn} \leq 5 \text{ minutes,} \\ 50 & \text{if } 5 < \text{dist}_{pn} \leq 15 \text{ minutes,} \\ 10 & \text{otherwise.} \end{cases}$$

If $\text{dist}_{pn} \leq 5$ minutes, the nursing home is in the same neighborhood, which is preferred by most patients. A drive between 5 and 15 minutes can still be seen as close by, whereas more than 15 minutes driving is rather far. We also developed corresponding waiting-time utilities based on the presumed preferences of patients in Amsterdam. We use the same groups FP and PP as for the small setting, i.e. the same interpretation but different parameter values. For the FP group, we have for both placement and preferred placement a linear waiting time utility, namely $v^U_{pn}(w_p) = 0.1w_p \ \forall \ p \in P^{FP}$ and $v^F_{pn}(w_p) = 0.1w_p \ \forall \ p \in P^{FP}$. For the PP group, we have that the waiting time utility to be placed in a preferred nursing home increases slightly faster, thus $v^F_{pn}(w_p) = 0.15w_p \ \forall \ p \in P^{PP}$.

For the placement waiting time utility, we developed a sigmoidal function such that the preferences of this group are best described according to elderly care experts. This function equals $v^U_{pn}(w_p) = \frac{100}{1+e^{-(0.09w_p-13)}} - 101 \; \forall \; p \in P^{PP}$, such that in combination with $g_{pn} \; \forall \; p \in P, n \in N$ we have at $w_p = 0$, all resulting utilities are negative except those of the preferred nursing homes. Then, after approximately three months, the utility becomes positive for the nursing homes with $\text{dist}_{pn} \leq 5$ minutes, after approximately 4.5 months, the utility becomes positive for the nursing homes with $\text{dist}_{pn} \leq 15$ minutes as well, and after approximately 6 months, the utility becomes positive for all (resulting) nursing homes.

Preferred nursing homes for patients are selected based on the highest utilities, with priority given to those with the highest scores. In cases where multiple nursing homes have equal utility, a random selection is made. To ensure fairness in comparing the model to the current situation, we selected preferred nursing homes for the current policy using the same utility scheme. The parameter values for the case study can be found in Table 6.2.

Parameter	Value	Parameter	Value
$\lvert N \rvert$	39	$v^U_{pn}(w_p)$	$0.1w_p \; \forall \; p \in P^{FP}$
λ	1.25/day	$v^F_{pn}(w_p)$	$0.1w_p \; \forall \; p \in P^{FP}$
θ^{-1}	666 days ($\approx$ 2 years)	$v^U_{pn}(w_p)$	$\frac{100}{1+e^{-(0.09w_p-13)}} - 101 \; \forall \; p \in P^{PP}$
μ^{-1}	666 days ($\approx$ 2 years)	$v^F_{pn}(w_p)$	$0.15w_p \; \forall \; p \in P^{PP}$
M	1000	c_n	Initial value: $20 \; \forall \; n \in N$
Δ	1 day		
g_{pn}	$\begin{cases} 100 & \text{if dist}_{pn} \leq 5 \text{ min} \\ 50 & \text{if } 5 < \text{dist}_{pn} \leq 15 \text{ min} \\ 10 & \text{else} \end{cases}$		

Table 6.2: Parameter values for case study Amsterdam.

6.3.4 Results

We analyzed various policies to assess their impact on performance measures. The results comparing the allocation model to the two extremes, as defined in Chapter 5, and the current policy are summarized in Table 6.3. Abandonment rates range from 6.0% for the shared queue to 36.9% for separate queues, indicating significant differences. Notably, under the separate queue policy, only 63.1% of individuals end up in their preferred nursing home, highlighting diverse bed demand across facilities in Amsterdam.

Table 6.3 illustrates the overall favorable performance of the allocation model. Compared to the current policy, abandonment rates decrease significantly from 32.2% to 7.4%, with the allocation model showing only slightly higher abandonment rates than the shared queue (which sets a lower bound). In addition, mean waiting times until placement and preferred placement are significantly reduced. However, the number of replacements increases in the allocation model compared to the current policy. Observe that the performance of the FP

Policy	% abandonments	% died at temp. NH	% died at pref. NH	MWT till placement (d)	MWT till preferred (d)	Mean nr. of replacements	Mean queue length	Mean occupancy
Shared queue	6.0% (5.3%-6.7%)	0.0% (0.0%-0.0%)	91.3% (87.4%-95.2%)	40 (36-43)	40 (36-43)	0.91 (0.87-0.95)	91 (82-100)	1.0 (1.0-1.0)
AM: tot	7.4% (7.2%-7.7%)	25.1% (24.8%-25.3%)	67.5% (67.2%-67.8%)	47 (46-48)	191 (189-193)	1.34 (1.33-1.34)	85 (83-87)	0.99 (0.99-0.99)
AM: FP	0.4% (0.4%-0.5%)	33.1% (32.8%-33.4%)	66.4% (66.1%-66.7%)	3 (2-3)	196 (192-199)	1.58 (1.58-1.58)		
AM: PP	14.5% (14.0%-14.9%)	16.9% (16.6%-17.2%)	68.6% (68.2%-69.0%)	99 (97-101)	186 (184-188)	1.09 (1.09-1.1)		
CP	32.2% (32.0%-32.4%)	4.5% (4.3%-4.6%)	63.3% (63.1%-63.6%)	234 (232-235)	257 (255-259)	0.86 (0.86-0.87)	313 (311-315)	0.72 (0.72-0.73)
Separate queues	36.9% (36.6%-37.1%)	0.0% (0.0%-0.0%)	63.1% (62.9%-63.4%)	256 (254-258)	256 (254-258)	0.63 (0.63-0.63)	350 (348-353)	0.67 (0.67-0.68)

Table 6.3: Results for the allocation model Amsterdam.
(Note: The values in parentheses are the 95% confidence intervals.)

and PP groups is well in line with their goals in terms of time to placement and preference.

Scenario of multiple preferred nursing homes
The results of the allocation model, applied to Amsterdam, when patients have two or three preferences are presented in Table 6.4. As anticipated, allowing two preferences leads to notable improvements, such as a considerable reduction in mean waiting time until placement in a preferred nursing home (from 177 to 105 days). In addition, the percentage of patients who end up in their preferred nursing home increases from 71.2% to 82.5%, albeit with an increase in the mean number of replacements. When the number of preferred nursing homes is increased to three, overall performance improves further, although not as dramatically as observed between one and two preferences. Therefore, the decision to adopt multiple preferences depends on balancing these improvements against the loss of specific nursing home preferences, particularly given that bed efficiency approaches its maximum level with two preferences (mean occupancy of 0.99). Healthcare managers must carefully consider these factors to determine the optimal strategy.

The impact of multiple preferred nursing homes within the current policy setting is explored in Table 6.5. These results reveal significant reductions in abandonment rates and mean waiting times when the current policy allows for more preferred nursing homes. However, compared to the allocation model, the current policy's performance is inferior even when

Policy	% abandonments	% died at temp. NH	% died at pref. NH	MWT till placement (d)	MWT till preferred (d)	Mean nr. of replacements	Mean queue length	Mean occupancy
AM(2): tot	6.6% (5.7%-7.5%)	8.5% (8.0%-9.1%)	84.8% (84.0%-85.6%)	40 (37-44)	98 (93-102)	1.35 (1.33-1.36)	75 (66-84)	0.99 (0.99-1.0)
AM(2): FP	1.1% (0.5%-1.6%)	14.6% (13.7%-15.4%)	84.3% (83.6%-85.1%)	6 (4-9)	102 (98-105)	1.69 (1.67-1.7)		
AM(2): PP	12.2% (10.9%-13.5%)	2.5% (2.1%-3.0%)	85.3% (84.1%-86.5%)	79 (73-85)	94 (88-100)	1.0 (0.98-1.03)		
AM(3): tot	6.3% (5.3%-7.3%)	2.5% (2.1%-2.9%)	91.2% (90.3%-92.1%)	43 (39-47)	61 (58-64)	1.27 (1.25-1.3)	86 (75-97)	1.0 (1.0-1.0)
AM(3): FP	4.2% (2.9%-5.4%)	5.0% (4.2%-5.9%)	90.8% (89.7%-91.9%)	26 (21-31)	61 (58-64)	1.63 (1.59-1.67)		
AM(3): PP	8.5% (7.6%-9.4%)	0.0% (0.0%-0.0%)	91.5% (90.6%-92.4%)	60 (57-64)	60 (57-64)	0.92 (0.91-0.93)		

Table 6.4: Results for the allocation model with multiple preferences in Amsterdam. (Note: The values in parentheses are the 95% confidence intervals.)

allowing the same number of preferred nursing homes. For example, with two nursing homes, the abandonment fraction is 14.7% for the current policy (CP(2)) and only 6.6% for the allocation model (AM(2)). Thus, the allocation model significantly outperforms the simple alternative of allowing patients to choose more than one preferred nursing home.

In summary, we find that if the system allows more patient replacements, the allocation model outperforms the current policy, significantly reducing waiting times in Amsterdam while also accommodating individual patient preferences.

Applied to a population of somatic patients in Amsterdam, the allocation model significantly reduced the abandonment fraction from 32.2% to 7.4% and decreased the mean time until placement by five months. However, these improvements required a higher rate of patient replacements, increasing from 0.86 to 1.34 on average. Furthermore, allowing patients to choose two preferred nursing homes further improved the model's performance, reducing the abandonment fraction to 6.6%. These findings underscore the relevance of the model in addressing the pressing need for efficient and sustainable long-term care solutions.

Policy	% abandonments	% died at temp. NH	% died at pref. NH	MWT till placement (d)	MWT till preferred (d)	Mean nr. of replacements	Mean queue length	Mean occupancy
CP(1)	32.2% (32.0%-32.4%)	4.5% (4.3%-4.6%)	63.3% (63.1%-63.6%)	234 (232-235)	257 (255-259)	0.86 (0.86-0.87)	313 (311-315)	0.72 (0.72-0.73)
CP(2)	14.7% (14.4%-15.0%)	0.0% (0.0%-0.0%)	85.3% (85.0%-85.6%)	100 (98-102)	100 (98-102)	0.85 (0.85-0.86)	164 (161-166)	0.91 (0.91-0.91)
CP(3)	7.5% (6.6%-8.4%)	0.0% (0.0%-0.0%)	92.5% (91.6%-93.4%)	51 (47-56)	51 (47-56)	0.92 (0.92-0.93)	100 (89-111)	0.98 (0.98-0.99)
CP(4)	6.8% (5.8%-7.7%)	0.0% (0.0%-0.0%)	93.2% (92.3%-94.2%)	46 (41-51)	46 (41-51)	0.93 (0.92-0.94)	94 (83-106)	0.99 (0.99-0.99)
CP(5)	6.7% (6.4%-7.0%)	0.0% (0.0%-0.0%)	93.3% (93.0%-93.6%)	45 (44-47)	45 (44-47)	0.93 (0.93-0.94)	97 (93-101)	0.99 (0.99-1.0)

Table 6.5: Current policy multiple preferences Amsterdam
Note: The values in parentheses are the 95% confidence intervals.

6.3.5 Scenario with popular nursing homes

For the region of Amsterdam, we also consider a scenario with more popular nursing homes. More specifically, in each of the eight parts of the city (called 'stadsdelen'), we randomly select a really popular nursing home. That is, with probability $p_{pop} = 0.25$, this nursing home is selected to be the preferred nursing home of a patient who belongs to that part of the city. The parts of the city and the popular nursing homes chosen are provided in Figure 6.9.

Table 6.6 shows the results for both the current policy and the allocation model. The table illustrates that in a scenario with popular nursing homes, the number of abandonments increases for both policies (recall that the abandonment fractions were 7.4% and 32.2% for the allocation model and the current policy, respectively). However, we see that compared to the situation without popular nursing homes, the percentage of abandonments under the allocation model only increases by 1.3%, while the percentage of abandonments under the current policy increases by 7.2%. This implies that the current policy is considerably more sensitive to a change in the popularity of nursing homes than the allocation model.

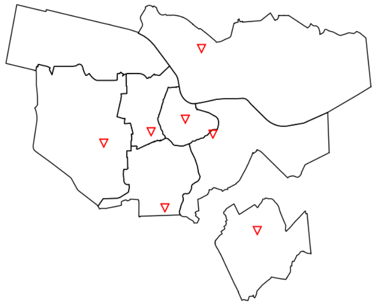


Figure 6.9: Selected popular nursing homes in city parts of Amsterdam.
(Note: In the city part in the north-west, Westpoort, no popular nursing home is selected, since no nursing home is located in this part.)

Policy	% Abandonments	% Died at temp. NH	% Died at pref. NH
AM: tot	8.6% (7.7%-9.4%)	72.5% (71.5%-73.5%)	18.7% (18.4%-19.2%)
AM: FP	0.7% (0.4%-0.9%)	79.2% (78.3%-80.1%)	20.1% (19.3%-21.1%)
AM: PP	16.3% (14.7%-17.7%)	66.3% (64.4%-68.0%)	17.5% (16.6%-18.5%)
CP	37.4% (36.7%-38.1%)	42.0% (41.0%-43.2%)	20.5% (19.4%-21.6%)

Table 6.6: Results scenario popular in Amsterdam.
(Note: The values in parentheses are the 95% confidence interval.)

6.4 Case Study: Rotterdam and Twente

6.4.1 Context and Data

In this section, we focus on two other different regions in the Netherlands: Rotterdam, a densely populated city region, and Twente, a rural area. The locations of the nursing homes were obtained from the website www.zorgkaartnederland.nl, which is an open database containing all nursing homes in the Netherlands. An extra check for address and care provided was done by inspection of the nursing homes' websites. Data on access requests for nursing homes and length-of-stays were obtained by non-public microdata from Statistics Netherlands. The data that were used as input for the model are summarized in Tables 6.7 and 6.8.

For both Rotterdam and Twente, a map is presented in Figure 6.10, where each triangle represents a nursing home in which older adults in need of somatic care reside. As can be seen, Rotterdam is an urban area with many nursing homes (39 in total) over a small area. In contrast, Twente is a rural region, with the nursing homes (62 in total) spread out over the geographical area. Public data by Statistics Netherlands indicate that both regions have approximately 600,000 inhabitants, whereas the surface of Twente is approximately

Parameter	Value	Parameter	Value
$\|N\|$ λ θ^{-1} μ^{-1}	39 1.11/day 666 days($\approx$ 2 years) 666 days($\approx$ 2 years)	$v_p^U(w_p)$	$\begin{cases} 0.1w_p & \forall\ p \in P^{FP1} \\ 0.1w_p - 11000 & \forall\ p \in P^{PP1} \\ w_p & \forall\ p \in P^{FP2} \\ w_p - 11000 & \forall\ p \in P^{PP2} \end{cases}$
M Δ	1000 1 day	$v_p^F(w_p)$	$\begin{cases} 0.1w_p & \forall\ p \in P^{FP1} \cup P^{FP2} \\ w_p & \forall\ p \in P^{PP1} \cup P^{PP2} \end{cases}$
g_{pn}	$\begin{cases} 100 & \text{if dist}_{pn} \leq 10 \\ 50 & \text{if } 10 < \text{dist}_{pn} \leq 20 \\ 10 & \text{else} \end{cases}$	c_n	Initial values: 50% 23 and 50% 24*

Table 6.7: Parameter values for case study of Rotterdam.

Parameter	Value	Parameter	Value
$\|N\|$ λ θ^{-1} μ^{-1}	62 1.46/day 666 days($\approx$ 2 years) 666 days($\approx$ 2 years)	$v_p^U(w_p)$	$\begin{cases} 0.1w_p & \forall\ p \in P^{FP1} \\ 0.1w_p - 11000 & \forall\ p \in P^{PP1} \\ 5w_p & \forall\ p \in P^{FP2} \\ 5w_p - 11000 & \forall\ p \in P^{PP2} \end{cases}$
M Δ	1000 1 day	$v_p^F(w_p)$	$\begin{cases} 0.1w_p & \forall\ p \in P^{FP1} \cup P^{FP2} \\ 5w_p & \forall\ p \in P^{PP1} \cup P^{PP2} \end{cases}$
g_{pn}	$\begin{cases} 100 & \text{if dist}_{pn} \leq 20 \\ 50 & \text{if } 20 < \text{dist}_{pn} \leq 30 \\ 10 & \text{else} \end{cases}$	c_n	Initial values: 50% 15 and 50% 16*

Table 6.8: Parameter values case study Twente

five times larger.

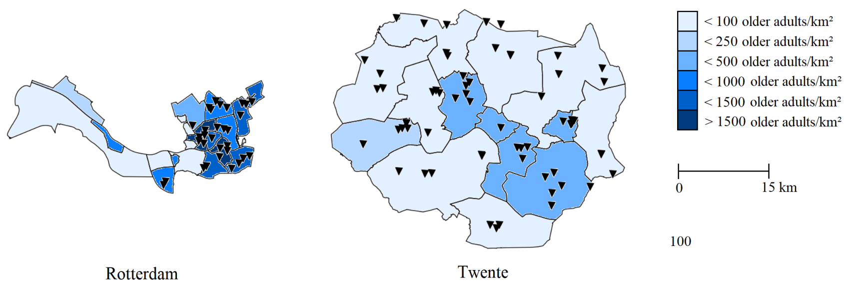


Figure 6.10: Focus regions with nursing homes and density of older adults (60+).

6.4.2 Results

The outcomes of the simulation studies are displayed in Table 6.9, where each output column corresponds to one of the selected outcome measures. We only report differences between values for which the 95%-confidence intervals are non-overlapping.

Metric	Rotterdam: Current practice	Rotterdam: Allocation model	Twente: Current practice	Twente: Allocation model
Mean waiting time (days) until placement	166 (162-169)	80 (75-85)	178 (175-181)	82 (66-98)
Mean waiting time (days) until pref. placement	544 (431-657)	315 (285-345)	637 (435-840)	265 (189-341)
Fraction of older adults that depart system from home	22.7% (20.5%-24.9%)	11.9% (10.8%-13.1%)	26.6% (22.0%-31.2%)	19.3% (17.4%-21.2%)
Fraction of older adults that depart system from temp. nursing home	43.8% (40.7%-46.9%)	38.8% (37.9%-39.6%)	51.3% (46.1%-56.5%)	42.3% (40.9%-43.6%)
Fraction of older adults that depart system from pref. nursing home	33.5% (30.8%-36.1%)	49.3% (47.4%-51.3%)	22.1% (19.9%-24.3%)	38.3% (35.3%-41.3%)

Table 6.9: Mean waiting times and departure fractions by policy and region.

In Table 6.9, we see that the allocation model outperforms the current practice policy for all output measures. First, the average waiting times show improvements: e.g., for Twente both the waiting time until placement and until preferred placement halve when the allocation model is used (from 178 and 637 days to 82 and 256 days, respectively). Next to that, more older adults eventually enter their preferred nursing home; for example, in Rotterdam this is 33.5% for the current practice policy and 49.3% for the allocation model. Hence, the results show that for both a rural region and an urban region, the allocation model is able to increase the efficiency in the system compared to the current practice.

Now consider the waiting times of the different preference profiles, as elaborated in Table 6.1. In Table 6.10, it can be seen that the Fast Placement preference profiles are placed within one day in a nursing home, which is in line with their preferences. Moreover, the quick placement is at the cost of obtaining a bed in a preferred nursing home, since the waiting time for a preferred placement is rather long. In contrast, the Preferred Placement profiles wait on average at least four months for a placement in a nursing home, but are placed faster into the nursing home of their preference. If an older adult chooses a Preferred Placement profile, the chance that this older adult ends up in a preferred nursing home is over 70% in Rotterdam and over 50% in Twente, which is above the current practice fractions of 33.5% and 22.1%, respectively (as provided in Table 6.9).

Furthermore, the profiles with two preference nursing homes are placed faster in preferred nursing homes than the profiles that selected one preferred nursing home. For example, for Twente, PP1 is placed in nine months (270 days) in a preferred nursing home, which reduces to six months (177 days) for PP2.

Metric	Rotterdam: FP1	Rotterdam: PP1	Rotterdam: FP2	Rotterdam: PP2
Mean waiting time (days) until placement	0 (0-1)	179 (167-191)	1 (0-1)	140 (132-148)
Mean waiting time (days) until pref. placement	515 (458-572)	179 (167-191)	425 (382-469)	140 (132-148)
Fraction of older adults that depart system from home	0% (0%-0%)	26.6% (24.7%-28.7%)	0% (0%-0%)	20.9% (18.6%-23.3%)
Fraction of older adults that depart system from temp. nursing home	80.7% (79.0%-82.5%)	0% (0%-0%)	74.5% (72.5%-76.5%)	0% (0%-0%)
Fraction of older adults that depart system from pref. nursing home	19.2% (17.4%-21.0%)	73.4% (71.3%-75.5%)	25.5% (24.2%-26.8%)	79.1% (76.7%-81.4%)

Metric	Twente: FP1	Twente: PP1	Twente: FP2	Twente: PP2
Mean waiting time (days) until placement	0 (0-1)	270 (208-332)	1 (0-1)	177 (136-218)
Mean waiting time (days) until pref. placement	418 (244-592)	270 (208-332)	386 (243-529)	177 (136-218)
Fraction of older adults that depart system from home	0% (0%-0%)	46.2% (42.1%-50.3%)	0% (0%-0%)	30.8% (28.1%-33.5%)
Fraction of older adults that depart system from temp. nursing home	88.5% (85.6%-91.4%)	0% (0%-0%)	80.9% (78.2%-83.6%)	0% (0%-0%)
Fraction of older adults that depart system from pref. nursing home	11.5% (9.4%-13.6%)	53.8% (50.7%-56.9%)	18.9% (15.7%-22.1%)	69.2% (65.3%-73.1%)

Table 6.10: Mean waiting times and departure fractions by preference profile and region.

6.5 Discussion

In this chapter, it was discussed how an allocation model can be implemented in practice using preference profiles. The performance of the allocation model was evaluated using simulation in three real-life case studies in the Netherlands: two urban areas (Amsterdam and Rotterdam) and a rural area (Twente). In comparison to the current practice policy, the allocation model reduced the waiting times until placement by at least a factor of two: from 313 to 85 days in Amsterdam, from 166 to 80 days in Rotterdam, and from 178 to

82 days in Twente. Moreover, the preferences of the older adults were also better retained: more older adults ended up in the nursing home of their desire and older adults that wanted to be placed fast were able to be placed within one day. Hence, these results show that the allocation model improved the waiting process in both efficiency and individual preferences.

Despite the promising results, with the implementation of the allocation model there are also organizational challenges and costs involved. As discussed in this chapter, regional collaboration is needed between a newly set-up placement office and the nursing homes. Moreover, the digital infrastructure for the database containing the available capacities and preferences must be developed, which is likely to be costly. Thus, due to the innovative nature of the allocation model, for the implementation an investment in both time and costs is required.

Besides the practical potential of the model, the allocation model presented in this chapter adds greater depth to the currently existing research on long-term care waiting-list management. A meta-review of waiting-list management specifically for the long-term care by Chafe, Coyte and Sears [42] showed that the majority of research on waiting-list management is focused on the organization of the waiting process, i.e., how waiting older adults can be taken care of. Little research has been performed on the organization of the waiting list, i.e., which older adult should be placed where, as a meta-analysis only reports two studies on that topic. Both those studies are about changing the priority setting from first-come-first-served to a needs-based criterion, so no mathematical model was involved [34, 116]. In view of that background, the allocation model based on preference profiles is a new concept in the geriatrics domain.

The care domain for older adults offers a wide spectrum of problems that can be improved or optimized by mathematical models. For example, studies have been conducted on optimal staffing strategies in nursing homes, shortest routes for staff members in home care services, and treatment scheduling for rehabilitation patients [18, 60, 153]. However, compared to mathematical research applied to other healthcare areas, such as the emergency departments or surgery planning, an unbalanced small fraction of research is devoted to long-term care [138]. Therefore, this research also contributes by broadening the scope of mathematics in the care for older adults setting.

6.6 Conclusions and Implications

The simulations for all case studies showed that the allocation model is a useful tool for reducing waiting times in care for older adults. These reduced waiting times without access to appropriate care can lead to reduced incidents [16]. Especially when an incident leads to hospitalization, high costs and hospitalization-associated disabilities are involved [141]. Therefore, the allocation model not only provides a solution for a reduction of waiting times directly (and the direct consequences such as reduced anxiety of older adults) [82], the allocation model might indirectly reduce demand for care in a broader sense as well.

The allocation model also offers opportunities other than allocating older adults to nursing homes. Namely, also other areas in which a scarcity of supply exists may benefit from the model, such as assigning older adults to home care or daycare. This is, however, slightly more complicated, since then ‘capacity’ is not defined as the number of beds, but as available hours per day. Hence, if the allocation model is extended to include this type of capacity as well, the allocation model can be used in other contexts within the care domain for older adults. Finally, the allocation model can also be extended and adapted in order to fit other healthcare contexts. We develop an allocation model for the psychiatric care domain in the next chapter.

Finally, this research reveals the potential of mathematical models in the care domain for older adults. We have shown that mathematical models can be preference-centered and thus not solely focused on efficiency at the expense of customization. As far as we are concerned, correctly developed mathematical models increase the cost-effectiveness of care and are indispensable in retaining the care for older adults’ expenses at acceptable level. Therefore, we suggest that future research focuses on developing logistically efficient solutions that are easy to implement, in order to contribute to solving the complex puzzle regarding the aging of the population.

7

Allocation of Persons with Resources: Knapsack-Based Routing

Based on:

Arntzen, R. J., Bekker, R., & van der Mei, R. D. (2024a). Knapsack-based routing for mental health placements: A split-horizon approach. *In review*.

Abstract

Objective: Persons with severe mental illness who cannot live independently can make use of residential facilities. The process of matching patients with suitable facilities is intricate, as both restrictions and client preferences play an important role. Regular placement policies inadvertently prioritize easy-to-place clients over hard-to-place ones, resulting in disproportionately long waiting times for the latter.

Method: To address this issue, we propose the Knapsack-Based Routing (KBR) model, which aims to allocate clients to server pools efficiently, considering both resource constraints and preferences. For this problem, we develop the so-called split-horizon method, utilizing an integer quadratic programming approach that incorporates the immediate and long-term consequences of scheduling decisions.

Results: Extensive simulation experiments demonstrate that the split-horizon method outperforms benchmark instances in two special cases of KBR: (i) Skill-Based Routing (SBR) with overflow, and (ii) queues with resource constraints where strategic idleness is crucial. Additionally, we present a benchmark instance for KBR with an optimality gap of 2.3%. Applying our method to the allocation process for mental health facilities in Amsterdam, the Netherlands, we observe a significant improvement over current policies, resulting in more equitable waiting times for both easy-to-place and hard-to-place clients.

Conclusion: Robust placement strategies in queueing systems with resource constraints must incorporate strategic idleness. Our split-horizon method provides a robust scheduling policy that ensures balanced waiting times between different client classes. This method is widely applicable to service operations that involve routing with resource requirements, such as in call centers and healthcare settings.

7.1 Introduction

The World Health Organization estimates that globally more than one in ten people live with a mental disorder, and 25-35% of the population will require mental health services at some point in their lives [194]. Common mental disorders include anxiety disorder (3.8%), depression (3.4%), bipolar disorder (0.6%), schizophrenia (0.3%), and eating disorders (0.2%) [56]. Substance use disorders, often associated with mental health issues, have a prevalence of 1.4% for alcohol use disorder, and 0.9% for drug use disorder [56]. The severity of these disorders significantly impacts individuals' daily lives. Approximately 1.2% of the global population suffers from severe mental illness, which impairs their ability to work, live independently, and maintain relationships [194].

Individuals unable to live independently due to severe mental illness can benefit from community residential facilities [152]. These facilities assist residents with daily activities such as cooking, medication management, finances, and social interaction, aiming to improve their overall functioning [96]. Group homes, a specific type of residential setting for psychiatric patients, foster communal living among residents who share common facilities like kitchens and living rooms [25, 97]. Studies have shown that group homes reduce (re)hospitalization rates, enhance employment prospects [61, 65], and promote independent living, self-esteem, and social functioning [114, 183]. Compared to supportive apartments where residents live independently, group home residents report higher levels of social support and less social isolation [126].

Achieving a suitable match between clients and housing facilities is crucial for effective living arrangements [66]. First, preferences play a significant role in client placement decisions. A home serves as an extension of a person's identity, offering security and fostering a sense of belonging and attachment. Consequently, it must align with the client's wishes to ensure positive health outcomes [31, 51]. Second, a suitable match involves meeting clients' specific resource needs within the facility and ensuring compatibility with current residents for maintaining a functional social environment [32]. Stringent requirements can pose challenges for clients in finding suitable accommodations quickly [26]. A standard placement policy often prioritizes easy-to-place clients over hard-to-place ones, resulting in disproportionately long waiting times for the latter [104]. The challenge lies in developing a placement policy that distributes waiting times more equitably among all client types.

Thus, the matching process between clients and facilities involves balancing preferences and constraints [69]. This process can be likened to a queueing model with multiple resource-constrained server pools, where efficient *routing* and *packing* of clients are essential. We introduce the Knapsack-Based Routing (KBR) model to address the challenge of efficiently allocating clients to facilities based on their resource requirements and preferences. This model finds broad applicability in contexts involving routing decisions with resource constraints, such as customer contact centers with multiskilled agents or hospital wards with limited capacity due to staffing and equipment constraints [7, 105].

The main contributions of this chapter are:

- *Modeling.* We define the KBR model, in which customer classes are allocated to server pools under resource constraints and preferences. The KBR model is an extension of both routing models and queueing systems with resource requirements. Firstly, the model extends Skill-Based Routing (SBR) models with overflow by incorporating additional resource requirements. Secondly, it extends queueing models with Heterogeneous Service Requirements (HSR), as well as models involving Resource Collaboration and Multitasking (RCM), by incorporating a network of resources and complex routing decisions. To determine the optimal policy for the KBR model, we solve a Markov Decision Process (MDP) and discuss the critical policy elements. In essence, a good-performing policy must be parameter-sensitive and state-dependent, addressing customer prioritization, strategic idleness, and overflow, while effectively managing the complex interactions among these components.
- *Design of split-horizon method.* Due to the curse of dimensionality, designing a heuristic method capable of solving real-life problems is imperative. To address this challenge, we develop the split-horizon method. This procedure divides the lookahead horizon into two segments: the finite horizon, which uses a novel scheduling technique, and the subsequent infinite horizon, which employs a value function approach based on a class-load approximation tailored to the resource framework. We evaluate our split-horizon method against benchmark policies for both SBR with overflow and HSR. Our findings demonstrate that, first, our method outperforms state-of-the-art policies for SBR with overflow. Second, it surpasses the best-known policies for HSR in scenarios where strategic idleness is critical. Additionally, we design a benchmark instance for the KBR model and find that the split-horizon method achieves an optimality gap of 2.3%.
- *Application to the mental health domain.* Our model contributes to the operations research literature in mental healthcare, an underexplored area [129]. We enhance existing knowledge by incorporating critical elements in mental health modeling, such as the categorization of individuals based on their resource needs and a focus on reducing waiting times. Furthermore, we apply our split-horizon method to a case study in Amsterdam, the Netherlands. Our findings demonstrate that this method can significantly improve upon current placement policies for group homes by reducing both overall waiting times and achieving more balanced waiting times between easy-to-place and hard-to-place customers. Thus, our approach holds considerable promise for enhancing mental health services.

7.2 Literature

The Operations Research literature applied to psychiatric care is sparse, as highlighted by a recent review titled "Mind the Gap", which identified only 13 papers on this subject [129]. We discuss three relevant papers on admission control for mental healthcare services. In their work, Heiner, Wallace and Young [81] explore admission control strategies for mentally ill individuals within service delivery systems. They propose a heuristic for resource allocation and program evaluation, emphasizing efficiency, effectiveness, and equity considerations. Equity is particularly critical due to variations in individual resource needs, adding complexity to decision-making. J. Wolpert and E. R. Wolpert [193] investigate the assignment of psychiatric patients to a network of community mental healthcare services, classifying patients by diagnosis and symptoms and identifying preferred treatment locations. They formulate a linear programming model to maximize patient treatment outcomes under a budget constraint. Another model developed by Leff, Dada and Graves [103] allocates clustered patient groups to community mental health services, allowing patients to change clusters based on their condition's improvement or deterioration. The objective is to assign "service package options" to maximize social welfare while adhering to resource constraints. These studies underscore the importance of categorizing individuals based on their resource needs as a fundamental aspect of allocating mental health services efficiently.

While the literature on Operations Research applied to psychiatric care remains limited, there has been a recent resurgence in interest, with half of the research published within the last decade. Noorain, Paola Scaparra and Kotiadis [129] emphasize the urgent need for models that address timely access to mental healthcare services, a significant practical challenge. They argue that new models must incorporate uncertainty and stochasticity, crucial features in mental healthcare settings characterized by issues such as long waiting times [136] and appointment no-shows [192]. In this study, we develop a stochastic model for mental healthcare, focusing on waiting times as a critical outcome measure.

In addition to the existing Operations Research literature in mental health, our research connects to technical literature in queueing theory. Specifically, the KBR model extends three types of queueing models: SBR with overflow, HSR, and RCM. Figure 7.1 provides an overview of these models, which we discuss in detail in the following paragraphs.

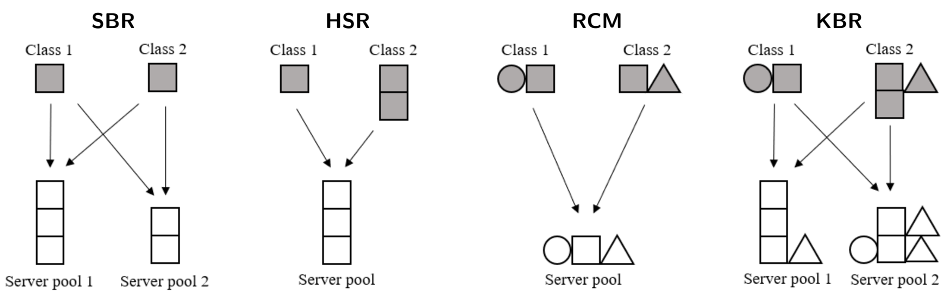


Figure 7.1: Schematic representation of related queuing models.

Due to the stochastic nature of our model and its allocation decisions, the KBR model can be seen as an extension of SBR models in queueing theory. SBR models involve matching customers of different classes to servers with specific skill sets [187]. Given the complexity of state and policy spaces in SBR models, exact analysis is often impractical. Therefore, researchers typically employ methods like heavy-traffic asymptotic analysis and heuristic policies. An important routing policy is the generalized $c\mu-$rule, which is defined as maximizing the instantaneous cost reduction rate. We define the queue length for customer class i as Q_i and a function for the waiting costs as $w_i(Q_i)$. Now, according to the generalized $c\mu-$rule, customer class i^* must be selected for queue j if $w'_{i^*}(Q_{i^*})\mu_{i^*j} \geq w'_i(Q_i)\mu_{ij}\ \forall i$. In a multi-class M/G/1 queue with linear holding cost, i.e. $w_i(Q_i) = w_iQ_i$, the $c\mu$-rule was found to be optimal [50]. Another class of policies is maximum pressure policies, in which the customer class i^* must be selected for queue j if $w_{i^*}X_{i^*}\mu_{i^*j} \geq w_iX_i\mu_{ij}\ \forall i$, where X_i denote the number of class i customers in the system. It was found that the maximum pressure policy is throughput optimal [53].

In our model, routing decisions consider not only capacity constraints and holding costs but also client preferences, akin to recent developments in SBR models discussed by Chen, Dong and Shi [44]. Their model features a parallel service system where each customer class is assigned to a dedicated server pool. When a customer is routed to a non-dedicated server pool, overflow costs incur. The authors propose a 'look-ahead' policy that balances holding costs against overflow costs to optimize routing decisions. This policy performs best in scenarios with small-sized server pools, while a modified maximum pressure policy is more effective for larger server pools. SBR models with overflow have found application in hospital ward allocation, where decisions involve balancing the wait for a dedicated ward against placing patients in alternative wards [45, 54, 151].

In addition to the SBR literature, the KBR model also extends to queueing models incorporating HSR. One of the pioneering studies involving resource requirements is conducted by Green [72], where a queueing system is analyzed with customers requiring varying numbers of servers. They propose a policy prioritizing customers with the fewest server requirements. Reiman [142] later provides approximations for blocking probabilities in such systems. More recently, Altman, Jiménez and Koole [6] examine admission control in telecommunications systems where customers have specific resource demands, akin to the stochastic knapsack problem. For large-scale scenarios, they derive a fluid model and identify trunk reservation policies as optimal. Similarly, Koeleman, Bhulai and van Meersbergen [99] address resource-constrained queueing in home care settings, proposing a heuristic based on trunk reservation that demonstrates strong performance compared to optimal policies. Such resource-constrained queues find applications across diverse service domains, including firefighting and emergency surgery [71].

A recent contribution to this field is provided by Zychlinski, Chan and Dong [202], who investigate a HSR queueing model. Customer class i may require multiple servers m_i simultaneously during service. The researchers develop the idle-avoid $\frac{c_i\mu_i}{m_i}$-rule, for which they show optimality for the preemptive case. For the non-preemptive case, the researchers develop the idle-aware policy. This rule is an optimization program to balance filling servers with class i customers with the highest instantaneous cost reduction rate and to reduce idleness by a Γ term, where the selection of Γ determines the index-based priority rule. Finally, the researchers note that for future research it would be "interesting to study a network of resources instead of a single type of resource. (...) The challenge then is to develop good scheduling policies that balance multiple resource constraints" (p. 35). This research answers the need for a non-preemptive scheduling policy for a system with multiple resource types.

Another relevant area of research pertains to models involving RCM. A typical example includes patient discharge processes, where the involvement of both a doctor and a nurse illustrates resource collaboration, and the doctor's multitasking requirement to be available at both patient diagnosis and discharge phases. Gurvich and Van Mieghem [76] highlight that such RCM networks inevitably experience idleness, indicating that even under maximum capacity, bottleneck resources cannot achieve full utilization. In deterministic settings, Bo, Dawande, Huh, Janakiraman and Nagarajan [28] explore methods to determine process capacities, while Dawande, Feng and Janakiraman [57] propose approaches to identify bottleneck resources through graph-theoretic analysis of process structures.

The study by Gurvich and Van Mieghem [77] investigates an RCM system akin to ours, specifically parallel M/M/1 queues with diverse resource requirements across different queues, similar to the model in Figure 7.1. They focus on prioritizing tasks that require extensive resource collaboration to optimize the system capacity. A hierarchical threshold priority policy is proposed to achieve more balanced queue sizes. Building on this, Özkan [135] extends the model by incorporating holding costs. The researcher devises a heuristic policy for the preemptive case that maximizes capacity utilization while minimizing holding costs. The adaptation to the non-preemptive case remains a challenge, as noted by the author who suggests that "deriving nonpreemptive control policies that perform well is not trivial and requires further research" (p. 301). Our KBR model can be conceptualized as an RCM model by setting resource coefficients as integers and limiting maximum capacity to one for each resource. Therefore, our allocation policy also provides a non-preemptive control policy applicable to RCM scenarios.

7.3 Model Formulation

The problem addressed in this chapter is how to route customers with heterogeneous service requirements to protected living facilities that we denote as server pools. We define I classes of customers and J server pools. The arrival process of a class i, $i = 1, \ldots, I$, customer entering the system is Poisson with rate λ_i. Moreover, the service time distribution of class i customers depends on server pool j, $j = 1, \ldots, J$, and is assumed to be exponential with rate $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{iJ})$. Each server pool j has capacity restrictions based on the availability of the N resources. We define the vector of available resources at server pool j as $\boldsymbol{b}_j = (b_{j1}, \ldots, b_{jN})$. Typically, the first resource $(n = 1)$ will denote the number of servers, such that b_{j1} represents the number of servers available in server pool j. Each customer class i has an associated demand for resources, denoted $\boldsymbol{\psi}_i = (\psi_{i1}, \ldots, \psi_{iN})$. For a customer to enter service, it is required that all resource requirements are met. The problem can be interpreted as involving multiple knapsacks, each filled with various types of items, determined by routing decisions. A schematic representation of the model is depicted in Figure 7.2.

The performance of the allocation mechanism is assessed using two metrics: waiting time and adherence to customer preferences. First, if customer class i has to wait before entering service, holding costs h_i are incurred. Second, if a customer is routed to a server pool that

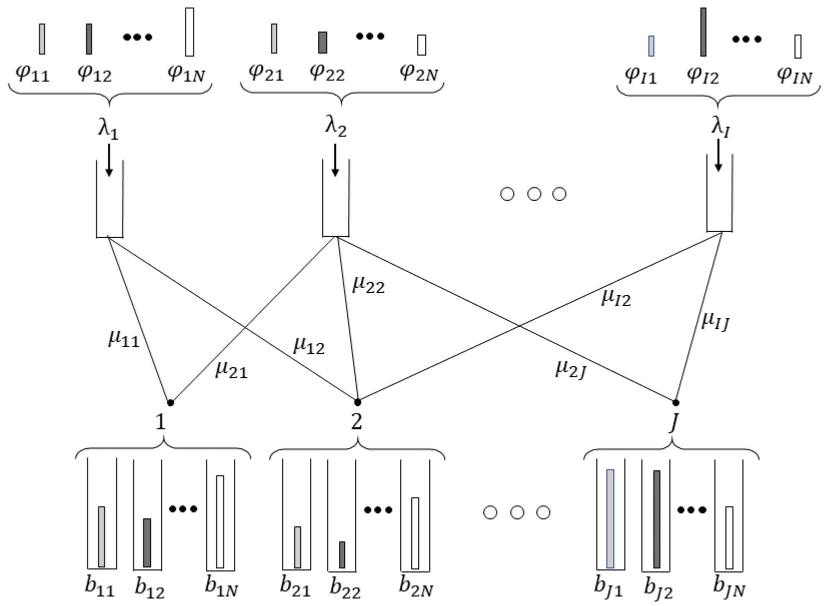


Figure 7.2: The KBR problem. Customer classes with demand for multiple resources need to be routed to one of the server pools with resource restrictions and class-dependent service times.

is not preferred, it represents a missed opportunity, which results in opportunity costs. The opportunity costs for class i are defined by $\boldsymbol{f}_i = (f_{i1}, \dots, f_{iJ})$.

Each class i customer is thus characterized by the tuple $(\lambda_i, \boldsymbol{\mu}_i, \boldsymbol{\psi}_i, h_i, \boldsymbol{f}_i)$. In some situations, it is useful to aggregate over customer classes that only differ in terms of their preferences, while they have identical parameters $\boldsymbol{\mu}_i$, ψ_i, and h_i. Let K be the number of such aggregate classes and let γ_{ik} be equal to 1 if customer class i belongs to aggregate class k, and 0 otherwise. More precisely, the map $i \to k$ is defined by $(\lambda_i, \boldsymbol{\mu}_i, \boldsymbol{\psi}_i, h_i, \boldsymbol{f}_i) \to (\bar{\lambda}_k, \boldsymbol{\mu}_k, \boldsymbol{\psi}_k, h_k)$, where $\bar{\lambda}_k = \sum_{i=1}^{I} \gamma_{ik}\lambda_i$, $\boldsymbol{\mu}_i = \boldsymbol{\mu}_k$, $\boldsymbol{\psi}_i = \boldsymbol{\psi}_k$, and $h_i = h_k$. Hence, class k can be interpreted as an aggregation over the classes i, ignoring their preferences.

We now describe the dynamics of the stochastic process. For the customers in the queue, their preferences are important, whereas preferences are no longer needed once the customer is taken into service. Denote the number of class $i \in \{1, \dots, I\}$ customers in the queue at time τ with $Q_i(\tau)$, and the number of class $k \in \{1, \dots, K\}$ customers being served in service pool j with $P_{kj}(\tau)$. These processes depend on the routing decisions. Let $A_{ij}(\tau)$ denote the total number of class i customers routed to service pool j by time τ. Moreover, let A_i and B_{kj} be unit rate Poisson processes modeling the arrival processes and service completion processes, respectively. Then, the system dynamics can be described as

$$
\begin{aligned}
Q_i(\tau) &= Q_i(0) + A_i(\lambda_i \tau) - \sum_{j=1}^{J} A_{ij}(\tau), \qquad i = 1, \ldots, I, \\
P_{kj}(\tau) &= P_{kj}(0) + \sum_{i=1}^{I} \gamma_{ik} A_{ij}(\tau) - B_{kj}\left(\mu_{kj} \int_0^{\tau} P_{kj}(t)\mathrm{d}t\right), \\
&\qquad k = 1, \ldots, K, \ j = 1, \ldots, J.
\end{aligned}
$$

The system state is a $(I + KJ)$-dimensional vector

$$
S(\tau) = \Big(Q_1(\tau), \ldots, Q_I(\tau), P_{11}(\tau), \ldots, P_{1J}(\tau), \ldots, P_{K1}(\tau), \ldots, P_{KJ}(\tau) \Big). \tag{7.1}
$$

We denote a given state by $s = (q_1, \ldots, q_I, p_{11}, \ldots, p_{1J}, \ldots, p_{K1}, \ldots, p_{KJ}) = S(\tau)$, and let $A(s)$ be the set of feasible actions, that is,

$$
A(s) = \Big\{ a \in \mathbb{N}^{I \times J} : a_{ij} \geq 0, \sum_{j=1}^{J} a_{ij} \leq q_i, \sum_{i=1}^{I} \psi_{ni} a_{ij} + \sum_{k=1}^{K} \psi_{nk} p_{kj} \leq b_{nj}, \ i = 1, \ldots, I, \ n = 1, \ldots, N, \ j = 1, \ldots, J \Big\}. \tag{7.2}
$$

The objective is to find a non-preemptive routing policy to control the allocation of customers to server pools, aiming to minimize the long-run average sum of holding and opportunity costs, defined as

$$
\lim_{\tau \to \infty} \frac{1}{\tau} \mathbb{E} \left[\sum_{i=1}^{I} \int_0^{\tau} h_i Q_i(t) \mathrm{d}t + \sum_{i=1}^{I} \sum_{j=1}^{J} f_{ij} A_{ij}(\tau) \right]. \tag{7.3}
$$

System stability

Due to the heterogeneity of the resource requirements, a condition for stability of the system cannot be expressed in terms of expected total demand and available capacity. In line with [202], we define a necessary condition for the vector of maximum arrival rates that can be processed under some allocation policy. To this end, we list all possible service configurations $x^1, \ldots, x^{S_j}$ in server pool j. In service configuration x^{s_j}, let $x_{ij}^s \in \mathbb{N}_0$ denote the number of class-i customers served in server pool j. The service configurations are constrained by resource availability, such that the set of feasible configurations for server pool j is defined by

$$\mathcal{X}_j = \left\{ x^{s_j} \in \mathbb{N}_0^I : \sum_{i=1}^{I} \psi_{in} x_{ij}^s \leq b_{jn}, \;\; n = 1, \ldots, N \right\}. \tag{7.4}$$

The maximum stability region can be characterized as

$$\mathcal{S} = \left\{ \boldsymbol{\lambda} \in \mathbb{R}_+^I : \lambda_i \leq \sum_{j=1}^{J} \sum_{s \in \mathcal{X}_j} \pi_{js} x_{ij}^s \mu_{ij}, \; \exists\, \pi_{js} \geq 0, \sum_{s \in \mathcal{X}_j} \pi_{js} \leq 1, \right. \tag{7.5}$$
$$\left. i = 1, \ldots, I, j = 1, \ldots, J \right\}.$$

In the definition above, π_{js} can be interpreted as the fraction of time that service configuration x^{s_j} is used in server pool j; the maximum service rate per time unit for class-i customers in server pool j available then reads $\sum_s \pi_{js} x_{ij}^s \mu_{ij}$.

7.4 Method

In this section, we start by considering small problem instances and find the optimal policy by numerically solving an MDP; see Appendix 7.A.1 for details on the MDP formulation. Our aim is to gain insight into, firstly, the optimal packing decisions and, secondly, the optimal routing decisions. Thereafter, we will illuminate our heuristic for the KBR model based on these insights.

7.4.1 Optimal packing strategy

For the packing problem, we seek to address questions such as: 'Is it optimal to develop a strict priority setting between the customer classes?' and 'Should we directly place a customer if there is capacity available?'. To answer these questions, we create an instance with one server pool, $J = 1$, one resource, $N = 1$, and two classes, $I = 2$. Classes 1 and 2 have resource coefficients $\psi_{11} = 1$ and $\psi_{21} = 2$, respectively. The availability of the resource is $b_{11} = 4$. This implies that we can place a maximum of four class 1 customers, two class 2 customers, or two class 1 customers and one class 2 customer. To visualize the policy, we consider an empty server pool, implying that the state is $(q_1, q_2, 0, 0)$. The occurrence of a particular state depends on the allocation policy; however, this is excluded from our current analysis. The actions (a_{11}, a_{21}) are defined by the number of customers of the two classes that are placed in the empty pool. We provide the action tables in Figure 7.3 for two different holding costs.

We find that for the case with holding costs $\boldsymbol{h} = (1, 3)$, we give strict priority to class 2, as $a_{21} = q_2$. This aligns with the $\frac{h_i \mu_{i1}}{m_i}$ rule from [202], shown to be optimal for the

$\boldsymbol{h} = (1, 3)$

q_2 \ q_1	0	1	2	3	4	5
2	(0,2)	(0,2)	(0,2)	(0,2)	(0,2)	(0,2)
1	(0,1)	(1,1)	(2,1)	(2,1)	(2,1)	(2,1)
0		(1,0)	(2,0)	(3,0)	(4,0)	(4,0)

$\boldsymbol{h} = (1, 1)$

q_2 \ q_1	0	1	2	3	4	5
2	(0,2)	(0,2)	(2,1)	(2,1)	(4,0)	(4,0)
1	(0,1)	(1,1)	(2,1)	(2,1)	(4,0)	(4,0)
0		(1,0)	(2,0)	(3,0)	(4,0)	(4,0)

Figure 7.3: Action tables (a_{11}, a_{21}) for $I = 2, J = 1, N = 1, \boldsymbol{b}_1 = (4), \boldsymbol{\psi}_1 = (1), \boldsymbol{\psi}_2 = (2), \mu_{ij} = 1, \lambda_i = 1, f_{ij} = 0, s = (q_1, q_2, 0, 0)$.

preemptive case. According to this rule, customer classes are prioritized by their highest instantaneous cost reduction rate, $\frac{h_i \mu_{i1}}{m_i}$, with m_i denoting the number of required servers. In our example, this rate is 1 for class 1 and 1.5 for class 2, favoring the placement of a class 2 customer over a class 1 customer.

For $\boldsymbol{h} = (1, 1)$, the instantaneous cost reduction rate for class 2 drops to 0.5, favoring class 1 over class 2. This is observed in Figure 7.3, where, e.g., state $(4, 2)$ results in action $(4, 0)$. However, placing a class 2 customer is prioritized over one class 1 customer in states like $(3, 2)$ with action $(2, 1)$ (instead of $(3, 0)$). Although the aggregate instantaneous cost reduction rate is 1 in both cases, the selected action minimizes idleness. Finally, note that the $\frac{h_i \mu_{i1}}{m_i}$-rule does not easily translate to our model with multiple resource requirements, lacking a clear 'm_i' calculation, and the rule is not optimal for the non-preemptive case [202]. Nevertheless, it provides valuable intuition on the trade-off between holding cost, service rate, and resource usage.

We now examine the impact of unbalanced holding costs, where one customer class becomes relatively more important. Setting the holding costs to $\boldsymbol{h} = (1, 5)$ and the arrival rates to $\lambda_i = 1.1$ and $\lambda_i = 1.3$, we present the corresponding action tables in Figure 7.4.

In Figure 7.4, for state $(1, 1)$, we observe action $(0, 1)$, indicating that the class 1 customer is not placed despite available resources. It appears optimal to keep two resource units idle for a future class 2 customer. This phenomenon is referred to as *strategic idleness*, where there is a deliberate decision not to place customers immediately, allowing for the placement of more challenging customers in the future. This concept is similar to the switching idleness discussed by Gurvich and Van Mieghem [77], where service capacity is intentionally kept

$\lambda_i = 1.1$

q_2 \ q_1	0	1	2	3	4	5
2	(0,2)	(0,2)	(0,2)	(0,2)	(0,2)	(0,2)
1	(0,1)	(0,1)	(2,1)	(2,1)	(2,1)	(2,1)
0		(1,0)	(2,0)	(3,0)	(4,0)	(4,0)

$\lambda_i = 1.3$

q_2 \ q_1	0	1	2	3	4	5
2	(0,2)	(0,2)	(0,2)	(0,2)	(0,2)	(0,2)
1	(0,1)	(0,1)	(0,1)	(0,1)	(0,1)	(2,1)
0		(1,0)	(2,0)	(2,0)	(4,0)	(4,0)

Figure 7.4: Action tables (a_{11}, a_{21}) for $I = 2, J = 1, N = 1, \boldsymbol{b}_1 = (4), \boldsymbol{\psi}_1 = (1), \boldsymbol{\psi}_2 = (2), \mu_{ij} = 1, \boldsymbol{h} = (1, 5), f_{ij} = 0, s = (q_1, q_2, 0, 0)$. The colored boxes contain actions in which there is strategic idleness.

idle to switch the class being served. However, strategic idleness is a more comprehensive approach, as it encompasses not only the switching of classes but also the adjustment of entire service configurations and the anticipation of future customer arrivals.

Moreover, increasing the arrival rates λ_i reduces the expected interarrival time of class 2 customers, leading to more states where strategic idleness is optimal. This happens according to a threshold structure in terms of the number of customers in the queue. For $\lambda_i = 1.1$, observe that strategic idleness is applied only in states like $(1, 1)$ but not in $(3, 0)$. This decision hinges on the time required for a class 2 customer to find an available place if all customers in the queue join the server pool, which is $\frac{1}{2\mu_{i1}} = \frac{1}{2}$ and $\frac{1}{3\mu_{i1}} = \frac{1}{3}$ time units, respectively. The former is deemed too long, yielding the preference for not placing the class 1 customer. Thus, actions are influenced by the types of classes occupying the server pool. For $\lambda_i = 1.3$, in state $(3, 0)$, the action becomes $(2, 0)$.

7.4.2 Optimal routing strategy

We now investigate the routing part. For this, we maintain the resource structure as introduced above but with two server pools ($J = 2$). Both class 1 and 2 prefer server pool 1, implying that $f_{i1} = 0$. If they are placed in server pool 2, the opportunity costs are $f_{i2} = 0.5$ or $f_{i2} = 2$. To evaluate the routing strategy, we fill server pool 1 with the maximum number of class 1 customers and leave server pool 2 empty, resulting in state $s = (q_1, q_2, 4, 0, 0, 0)$. The actions are now provided in Figure 7.5.

$f_{i2} = 2$

q_2						
2	(0,0)	(0,0)	(0,1)	(0,1)	(0,1)	(0,2)
1	(0,0)	(0,0)	(0,0)	(0,0)	(0,1)	(0,1)
0		(0,0)	(0,0)	(0,0)	(0,0)	(0,0)
q_1	0	1	2	3	4	5

$f_{i2} = 0.5$

q_2						
2	(0,2)	(0,2)	(0,2)	(0,2)	(2,1)	(2,1)
1	(0,1)	(0,1)	(0,1)	(2,1)	(2,1)	(2,1)
0		(0,0)	(1,0)	(2,0)	(3,0)	(4,0)
q_1	0	1	2	3	4	5

Figure 7.5: Action tables with actions $(a_{11}, a_{21}, a_{12}, a_{22})$ denoted by (a_{12}, a_{22}) as $a_{11} = a_{21} = 0$, for $I = 2, J = 2, N = 1, b_{jn} = 4, \boldsymbol{\psi}_1 = (1), \boldsymbol{\psi}_2 = (2), \mu_{ij} = 1, \lambda_i = 1, h_i = 1, \boldsymbol{f}_i = (0, f_{i2}), s = (q_1, q_2, 4, 0, 0, 0)$.

For the situation in Figure 7.5, it holds that $a_{11} = a_{21} = 0$ for all cases as server pool 1 is at full capacity. Moreover, overflowing is primarily determined by comparing the waiting time and corresponding holding costs in the preferred server pool against the opportunity costs. With holding costs set at $\boldsymbol{h} = (1, 1)$, we prioritize class 1 customers, resulting in longer wait times for class 2 customers, who are consequently placed in server pool 2 sooner. When $f_{i2} = 2$, only class 2 customers overflow. Here, we observe that the overflow decision depends on the interaction between parameter values and resource requirements. For $f_{i2} = 0.5$, we overflow both customer types, depending on the number of customers of each class in the queue. When $q_2 = 0$, a threshold structure applies to class 1. For instance, in state $(2, 0)$, the second class 1 customer in the queue is overflowed due to holding costs exceeding the overflow costs, with action $(0, 0, 1, 0)$. Adding another class 1 customer results in an increase of a_{12} by one; i.e., in state $(3, 0)$ the optimal action is $(0, 0, 2, 0)$.

7.4.3 Policy features

After evaluating the optimal policy for the KBR model, we have identified pivotal elements for an effective allocation policy. Typically, for multiclass queueing networks, index-based priority rules are employed [127]. However, due to the resource requirements, the KBR model necessitates deliberate idleness, as an index-based priority rule inadvertently allocates empty capacity to easy-to-place customers. Therefore, a robust policy requires *scheduling* to plan which customers to place immediately and which to defer, potentially integrating them into efficient service configurations.

As inferred from the optimal policy, the scheduling policy must be parameter-sensitive, considering the values for arrival rate, service rate, holding costs, opportunity costs, and resource requirements. Additionally, it must facilitate state-dependent decision-making. Specifically, the policy must determine (i) which customers to *prioritize* in the available capacity, (ii) which customers to *defer* to accommodate hard-to-place customers later, thereby applying strategic idleness, and (iii) which customers to *overflow*. Moreover, the policy must effectively navigate the complex interactions among these components (i)-(iii), particularly in a stochastic environment, to arrive at a good-performing decision. Finally, the policy should employ a monotone structure, resulting in a threshold policy for decision-making.

7.4.4 Split-horizon method

We have designed a method, formulated in Section 7.4.7, that meets our aims for a good-performing placement policy for the KBR model. The general idea is to use a rolling-horizon approach that is evoked at the time of an event. At time 0, we make a rough allocation plan by looking ahead. Although the plan looks ahead, we only perform allocations at time 0; the rest of the plan considers future costs. Upon the next event, we reset the time at $t = 0$ and make a new future plan, but only carry out the new allocation for $t = 0$.

The allocation planning is created by splitting the look-ahead horizon into two parts: the finite horizon $\{0, 1, \ldots, T\}$ and the infinite horizon beyond time T, as illustrated in Figure 7.6. The goal of the short-term planning in segment I is to accommodate for the complex decision process that results from the heterogeneous resource requirements. The purpose of the long-term planning in segment II is to appropriately balance between carrying out sub-optimal allocations and the burden on the system for postponing such allocations.

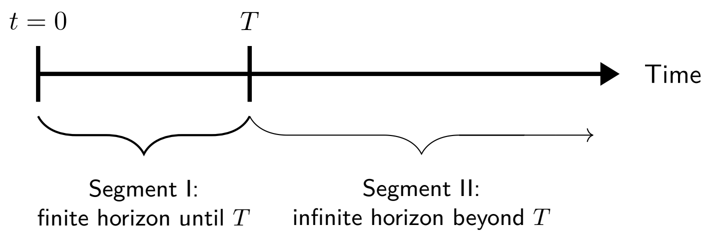


Figure 7.6: Split of the look-ahead horizon.

This composite approach is inspired by multi-step look-ahead strategies [24]. More specifically, for a given state $s_0 \in \mathcal{S}(0)$ at time 0, the aim is to achieve

$$\min_{a_0, \tilde{a}_1, \ldots, \tilde{a}_T} \mathbb{E}\left\{ c(s_0, a_0) + \sum_{k=1}^{T} c(s_k, \tilde{a}_k(s_k)) + V(s_T^*) \right\}, \tag{7.6}$$

where a_0 is the allocation at time 0, $\tilde{a}_k$ is the allocation policy at time k, $c(s, a)$ is the direct cost for allocation a in state s, and s_T^* is the post-action state at time T. Here, $V(\cdot)$ is the terminal cost, which we approximate using the relative value function of a base policy. If we set $T = 0$, the second term in Equation (7.6) vanishes, and the minimization involves only the direct cost of the action at $t = 0$ and the resulting future costs $V(s_0^*)$. In this case, the method can be interpreted as a one-step policy improvement.

The approximate cost of the first T steps (segment I) is addressed in Section 7.4.5, whereas the terminal cost (segment II) is discussed in Section 7.4.6. The minimization problem culminates in an Integer Quadratic Program presented in Section 7.4.7.

7.4.5 Segment I: finite horizon up to T

For the finite horizon up to time T, we develop a rough allocation plan. As this plan needs to be scalable, which is challenging due to the stochasticity combined with the large state and action spaces, it is based on the expected resource utilization. More specifically, let Δ be the time between decision epochs such that the sequence of time epochs can be described as $t_n = t_{n-1} + \Delta$, for $n = 1, 2, \ldots$. Denote the service time of a class i customer in server pool j by the random variable B_{ij} and the residual service time of class k in server pool j by B_{kj}^r. Consider the amount of resource n used at time t in server pool j, defined by $R_{jn}(t)$. Let the planned allocation be given by $x_{ijt} \in \mathbb{N}$ which denotes the number of class i clients allocated to j at time t. Then, the random variable $R_{jn}(t)$ can be described by the identity

$$R_{jn}(t) = \sum_{k=1}^{K} \psi_{kn} p_{kj} \mathbb{1}_{\{B_{kj}^r \geq t\}} + \sum_{i=1}^{I} \sum_{s=0}^{t} \psi_{in} x_{ijs} \mathbb{1}_{\{B_{ij} \geq t-s\}}. \tag{7.7}$$

For the rough planning procedure, we assume that the *expected* amount of resources should be sufficient, that is, $\mathbb{E}[R_{jn}(t)] \leq b_{jn}$. Since we assume exponential service times, it holds that $\mathbb{P}(B_{kj}^r \geq t) = e^{-\Delta \mu_{kj} t}$ and $\mathbb{P}(B_{ij} \geq t - s) = e^{-\Delta \mu_{ij}(t-s)}$. Clearly, the assumption of exponential service times is not essential, and the planning can be easily adapted to a non-exponential setting. Using Equation (7.7), we find

$$\mathbb{E}[R_{jn}(t)] = \sum_{k=1}^{K} \psi_{kn} p_{kj} e^{-\Delta \mu_{kj} t} + \sum_{i=1}^{I} \sum_{s=0}^{t} \psi_{in} x_{ijs} e^{-\Delta \mu_{ij}(t-s)}. \tag{7.8}$$

We also consider costs in the allocation planning. Firstly, if a class i customer is placed in server pool j, we incur an opportunity cost of f_{ij}. Additionally, we account for holding costs if a customer is placed in a server pool at a later time; at time t for class i, this results in a cost of $t\Delta h_i$. If a customer cannot be placed within the time horizon, we incur holding costs over this period, amounting to $T\Delta h_i$. These 'leftover' customers still impact the system beyond time T. Consequently, we also incur costs for these customers over the infinite horizon after time T, as further explained in the next section.

7.4.6 Segment II: infinite horizon beyond T

For the infinite horizon beyond T, we approximate the value function term $V(\cdot)$ in Equation (7.6), following the intuition of multi-step look-ahead strategies. As a base policy, we assign a dedicated capacity to each class, an approximation that is also used in the look-ahead policy of [44]. This is based on the premise that if a class i customer is not served during $t = 0, \dots, T$, the customer will wait to be served in a server pool j where $f_{ij} = 0$. Consequently, we need to determine the load of class i on the system, which is not directly derived from conventional equations due to the heterogeneous resource requirements. To characterize the class load for each class i customer, we first employ a notion of utilization similar to that used in, for example, [90].

Class load approximation

The load is typically a key ingredient in many performance measures in classical queueing systems. We fix $c_i = \lambda_i/\bar{\lambda}$, with $\bar{\lambda} = \sum_i \lambda_i$, as the fraction of arrivals of class i. Our approach is to increase the aggregate arrival rate $\bar{\lambda}$ until the boundary of the stability region is reached, using dedicated capacity only, from which we derive the corresponding utilization. By dedicated capacity, we mean that class i can only utilize server pool j under the condition that $f_{ij} = 0$. Let ω_{ij} be 1 if $f_{ij} = 0$ and ∞ otherwise. We are interested in the service configurations in server pool j that include only the dedicated classes, denoted by $\mathcal{X}_{j(\omega)}$, which are given by

$$\mathcal{X}_{j(\omega)} = \left\{ x^{s_j} \in \mathbb{N}_0^I : \sum_{i=1}^{I} \psi_{in} x_{ij}^{s} \omega_{ij} \leq b_{jn}, \;\; n = 1, \dots, N \right\}. \tag{7.9}$$

Moreover, we are interested in the set of *efficient* service configurations for server pool j, which is a subset of $\mathcal{X}_{j(\omega)}$. Our interest is in the Pareto efficient set $\mathcal{X}^E_{j(\omega)}$, in which it is not possible to serve an additional customer, i.e., no obvious service capacity is wasted. More specifically, this set is given by

$$\mathcal{X}^E_{j(\omega)} = \Bigg\{ x^{s_j} \in \mathcal{X}_{j(\omega)} \mid \nexists\, x^{s'_j} \in \mathcal{X}_{j(\omega)} \text{ such that } \forall i, \\ x_{ij}^{s'} \geq x_{ij}^{s} \text{ and } \exists i,\; x_{ij}^{s'} > x_{ij}^{s} \Bigg\}. \tag{7.10}$$

In line with the maximum stability region, we need to find the fraction of time π_{js} in which the efficient service configuration $x^{s_j} \in \mathcal{X}^E_{j(\omega)}$ is employed. Thus, our goal is to maximize the aggregate arrival rate $\bar{\lambda}$, retaining the proportions of arrivals per class, and to find the corresponding vector of time fractions π^*_{js} for each server pool j. This leads to the following optimization problem

$$\pi^*_{js} = \arg\max_{\pi_{js}} \Bigg\{ \bar{\lambda} : \lambda_i \leq \sum_{j=1}^{J} \sum_{s \in \mathcal{X}^E_{j(\omega)}} \pi_{js} x^s_{ij} \mu_{ij}, \ \lambda_i = c_i \bar{\lambda}, \tag{7.11}$$
$$\sum_{s \in \mathcal{X}^E_{j(\omega)}} \pi_{js} = 1, i = 1, \ldots, I, j = 1, \ldots, J \Bigg\}.$$

Using the solution of (7.11), we determine the maximum average service rate per time unit for class i, denoted as s^*_i, and approximate the corresponding class-i utilization as follows:

$$\rho_i = \frac{\lambda_i}{s^*_i}, \text{ with } s^*_i = \sum_{j=1}^{J} \sum_{s \in \mathcal{X}^E_{j(\omega)}} \pi^*_{js} x^s_{ij} \mu_{ij}. \tag{7.12}$$

We illustrate the procedure with a small healthcare example. A nurse can handle up to four class 1 patients, two class 2 patients, or two class 1 and one class 2 patients. Additionally, since class 2 patients require a physician, only one class 2 patient can be managed at a time. We find the corresponding efficient service configurations $\mathcal{X}^E_{1(\omega)} = \{(2,1),(4,0)\}$ as shown in Figure 7.7. Via the optimization problem in (7.11), we find potential maximum λ_i's equal to $\lambda_1 = 1.5$ and $\lambda_2 = 1$, and the efficient time fractions $\pi^*_{1s} = (0.8, 0.2)$. Finally, the maximum service rates per time unit are now calculated as in (7.12) and are equal to $s^*_1 = 2$ and $s^*_2 = 1$.

The class-load description introduced above is typically appropriate when customer classes have only a single preferred server pool. However, when customers have multiple preferred server pools, it is a widely shared phenomenon that, in queueing situations where some customer classes prefer multiple service pools, the system tends to behave approximately as if all classes can be routed to all service pools [11, 121, 187]. In other words, with a limited amount of flexibility, nearly all economies of scale can be achieved. This means that customer waiting times are better approximated by aggregating capacity across the service pools and disregarding the specific preferences of the customer classes, implying that $\omega_{ij} = 1$ for all i, j. Consequently, in this case, it is better to base the value function, which will be discussed in the next section, on the aggregate classes defined in Section 7.3.

Value function approximation

We now quantify the burden on the system for every waiting customer at time T using the value function $V(\cdot)$ in Equation (7.6). As a base policy, each class is assigned a dedicated

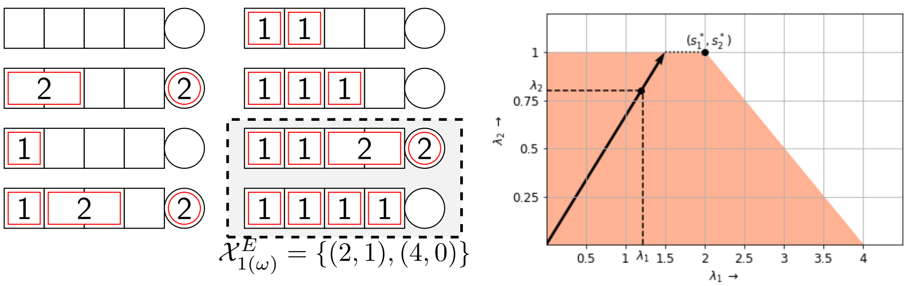


Figure 7.7: Example procedure for obtaining s_i^* for $I = 2, J = 1, N = 2, \boldsymbol{b}_1 = (4,1), \boldsymbol{\psi}_1 = (1,0), \boldsymbol{\psi}_2 = (2,1), \boldsymbol{\lambda} = (1.2, 0.8), \mu_{ij} = 1, f_{ij} = 0$. We depict $\mathcal{X}_{1(\omega)}$ in the left plot and find $\mathcal{X}^E_{1(\omega)} = \{(2,1),(4,0)\}$. Presented in the right picture is the stability region $(2,1)\pi^*_{11} + (4,0)\pi^*_{12}$, where $\pi^*_{11}, \pi^*_{12} \in [0,1]$ and $\pi^*_{11} + \pi^*_{12} \leq 1$. It can be seen that class 2 is binding as the class ratios yields a maximum of $\lambda_1 = 1.5$, while $s_1^* = 2$.

capacity. We choose the capacity of class i by taking s_i^*, as given by (7.12), which represents the maximum service rate per time unit assuming the fractions of arrivals per class remain the same. For a tractable functional form of the value function, we use the value function of an M/M/1 queue, which is $V(x) = x(x+1)/(2(\mu - \lambda))$, where λ denotes the arrival rate and μ the service rate. Hence, with y_i customers of class i waiting, we approximate the incurred holding costs by $h_i y_i^2/(2(s_i^* - \lambda_i))$. The quadratic term in y_i is both convenient for the optimization model and also accounts for the phenomenon of "waiting on waiting".

To provide some intuition, we note that the value function leads to a threshold-type overflow policy. Taking the derivative of $h_i y_i^2/(2(s_i^* - \lambda_i))$ with respect to y_i, we find that the $\bar{y}_i$-th waiting class i customer incurs an additional waiting cost of $h_i \bar{y}_i/(s_i^* - \lambda_i)$, which is a linear increasing function in $\bar{y}_i \in \mathbb{N}$. From Section 7.4.2 it follows that overflow occurs if the expected waiting cost for a customer exceeds the opportunity costs ϕ. Thus, the threshold for preferring overflow over waiting will approximately be

$$\arg\min \left\{ \bar{y}_i \mid \frac{h_i \bar{y}_i}{s_i^* - \lambda_i} > \phi \right\}. \tag{7.13}$$

7.4.7 Integer Quadratic Program

Combining Sections 7.4.5 and 7.4.6, we propose a first heuristic that is based on the customers currently in the system. At time 0, the action is to allocate x_{ij0} class i customers to j, which follows from the following Integer Quadratic Program (IQP) model:

$$\min \underbrace{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=0}^{T}(f_{ij}+t\Delta h_i)x_{ijt} + \sum_{i=1}^{I} h_i T\Delta\left(q_i - \sum_{j=1}^{J}\sum_{t=0}^{T} x_{ijt}\right)}_{\text{Segment I: finite horizon}}$$

$$+ \underbrace{\sum_{i=1}^{I} \frac{h_i\left(q_i - \sum_{j=1}^{J}\sum_{t=0}^{T} x_{ijt}\right)^2}{2(s_i^* - \lambda_i)}}_{\text{Segment II: infinite horizon}} \tag{7.14}$$

$$\text{s.t.} \quad \mathbb{E}[R_{jn}(t)] \le b_{jn} \qquad j=1,\ldots,J,\ n=1,\ldots,N,\ t=0,\ldots,T, \tag{7.15}$$

$$\sum_{j=1}^{J}\sum_{t=0}^{T} x_{ijt} \le q_i \qquad i=1,\ldots,I, \tag{7.16}$$

$$x_{ijt} \in \mathbb{N} \qquad i=1,\ldots,I,\ j=1,\ldots,J,\ t=0,\ldots,T. \tag{7.17}$$

where $\mathbb{E}[R_{jn}(t)]$ and s_i^* are given by (7.8) and (7.12), respectively.

The objective (7.14) consists of three parts. The first part corresponds to the opportunity and holding costs for customers allocated during $\{0,1,\ldots,T\}$. The second part depicts the holding cost during the time horizon for customers who are not scheduled yet. The third part represents the terminal cost of those customers, given by the value function. We refer to Remark 7.1 below in case aggregation of classes is appropriate. Constraint (7.15) ensures that the expected resource usage from (7.8) does not exceed resource availability at any time t. The number of class i customers that can be allocated to a server pool is restricted by the number of customers waiting in the corresponding queue, q_i, see (7.16). Finally, the integer constraints are given by (7.17).

Remark 7.1. In case customers have multiple preferred server pools, aggregating classes for the value function approximation seems desirable, see Section 7.4.6. The only modification required concerns the third term of the objective function, which represents segment II. Specifically, (7.14) should be replaced by

$$\sum_{k=1}^{K} \frac{h_k\left\{\sum_{i=1}^{I}\gamma_{ik}\left(q_i - \sum_{j=1}^{J}\sum_{t=0}^{T} x_{ijt}\right)\right\}^2}{2(s_k^* - \lambda_k)}.$$

Remark 7.2. Equation (7.8) shows that the expected resource use of a customer in service decreases over time but never entirely vanishes, as the service time has infinite support. This can complicate allocation plans, especially in systems with limited server capacity. To address this, we use a bound ξ: if $\mathbb{P}(B_{kj}^r \ge t) < \xi$, we set $\mathbb{P}(B_{kj}^r \ge t) = 0$.

IQP including future arrivals

Next, we aim to develop a second heuristic that incorporates the impact of future arrivals. Given the considerable variability in the future arrival process and the presence of multiple customer types, an exact approach is infeasible. Therefore, we employ a heuristic method. Specifically, we propose reserving slots in server pools for customer allocation. Let $\tilde{a}_{ijt} \in \mathbb{N}$ be the number of class i slots reserved at server pool j at time t, which will typically be 0 or occasionally 1. Define $\tilde{A}_{jn}(t)$ as the amount of resource n used in server pool j at time t by these reserved slots, assuming that they are utilized. We then obtain

$$\mathbb{E}[\tilde{A}_{jn}(t)] = \sum_{i=1}^{I}\sum_{s=0}^{t} \psi_{in}\tilde{a}_{ijs}e^{-\Delta\mu_{ij}(t-s)}. \tag{7.18}$$

For the rough scheduling, we require that the expected resource utilization due to both current customers and reserved slots be within the available capacity. This implies the constraint $\mathbb{E}[R_{jn}(t)] + \mathbb{E}[\tilde{A}_{jn}(t)] \leq b_{jn}$. Additionally, we need to estimate the cost associated with new customers, focusing exclusively on holding costs, which we refer to as backlog costs. Let $z_i(t)$ be the class i backlog at time t. Clearly, the class i backlog follows the recursion relation

$$z_i(t) = \max\left\{ z_i(t-1) + \tilde{A}_i(t-1,t) - \sum_{j=1}^{J}\tilde{a}_{ijt}, 0\right\},$$

with $\tilde{A}_i(t_0, t_1)$ the number of class i arrivals between epochs t_0 and t_1. In the heuristic, we use the expected value $\lambda_i\Delta$ as an approximation for $\tilde{A}_i(t-1,t)$. Moreover, the recursion can be made linear using standard arguments. Combining the above, we use the following

Mixed Integer Quadratic Program (MIQP) for the rough planning at time 0:

$$\min \underbrace{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=0}^{T}(f_{ij}+t\Delta h_i)x_{ijt}+\sum_{i=1}^{I}\sum_{t=0}^{T}h_i\Delta z_i(t)+\sum_{i=1}^{K}h_iT\Delta y_i}_{\text{Segment I: finite horizon}}$$

$$+ \underbrace{\sum_{i=1}^{K}\frac{h_i y_i^2}{2(s_i^*-\lambda_i)}}_{\text{Segment II: infinite horizon}} \tag{7.19}$$

$$\text{s.t. } \mathbb{E}[R_{jn}(t)]+\mathbb{E}[\tilde{A}_{jn}(t)]\leq b_{jn} \quad j=1,\ldots,J,\ n=1,\ldots,N,\ t=1,\ldots,T, \tag{7.20}$$

$$\sum_{j=1}^{J}\sum_{t=0}^{T}x_{ijt}\leq q_i \qquad i=1,\ldots,I, \tag{7.21}$$

$$z_i(t)\geq z_i(t-1)+\lambda_i\Delta-\sum_{j=1}^{J}\tilde{a}_{ijt} \qquad i=1,\ldots,I,\ t=0,\ldots,T, \tag{7.22}$$

$$y_i\geq q_i-\sum_{j=1}^{J}\sum_{t=0}^{T}x_{ijt}+z_i(T)-1 \qquad i=1,\ldots,I, \tag{7.23}$$

$$z_i(t)\geq 0 \qquad i=1,\ldots,I,\ t=0,\ldots,T, \tag{7.24}$$

$$x_{ijt}\in\mathbb{N} \qquad i=1,\ldots,I,\ j=1,\ldots,J,\ t=0,\ldots,T, \tag{7.25}$$

$$\tilde{a}_{ijt}\in\mathbb{N} \qquad i=1,\ldots,I,\ j=1,\ldots,J,\ t=0,\ldots,T. \tag{7.26}$$

In this second heuristic, we have that the objective is extended by the term containing the holding costs for the residual future arrivals in (7.19). Moreover, we have in (7.20) that the expected resource used also contains the potential placement of future arrivals. Constraint (7.21) ensures that the number of placed customers does not exceed the queue length. In (7.22), the backlog recursion of the future arrivals is modeled. In (7.23), we find that the residual class i customers y_i include the residual future arrivals $z_i(T)$. As $z_i(T)$ is a float and y_i is an integer, we subtract one from the right-hand side and use a greater or equal sign. Finally, the variables are defined in (7.24), (7.25), and (7.26).

In the following, if we denote the term 'split-horizon method', we refer to the method without taking into account future arrivals, thus the IQP model in (7.14)-(7.17). The IQP including future arrivals, presented in (7.19)-(7.26), will be denoted as split-horizon+ method.

7.5 Results

First, we compare the heuristic results to benchmark instances of SBR with overflow and HSR. Thereafter, we obtain for our benchmark instance the optimality gap. Lastly, the heuristic is applied to a real-life case study.

7.5.1 Comparison to SBR benchmark

The KBR model, with $N = 1$, simplifies to an SBR model with overflow. We evaluate the split-horizon method using the benchmark instance from Chen, Dong and Shi [44], adapted for discrete time. This instance involves three classes ($I = 3$) and three server pools ($J = 3$), with opportunity costs $\phi \in \{0.2, 2, 10\}$ penalizing non-dedicated assignments. Since $\phi = 0.2$ and $\phi = 2$ yield similar overflow strategies, we focus on $\phi = 2$ and $\phi = 10$, also varying arrival rates. The study examines several allocation policies, presenting results for the top three: the modified $c\mu$-rule, the modified maximum pressure policy, and the look-ahead policy. For the set-up of these policies, we refer to Chen, Dong and Shi [44]. For the split-horizon method, we set $T = 0$, as there are no resource requirements, eliminating the need for scheduling. The results are provided in Figure 7.8.

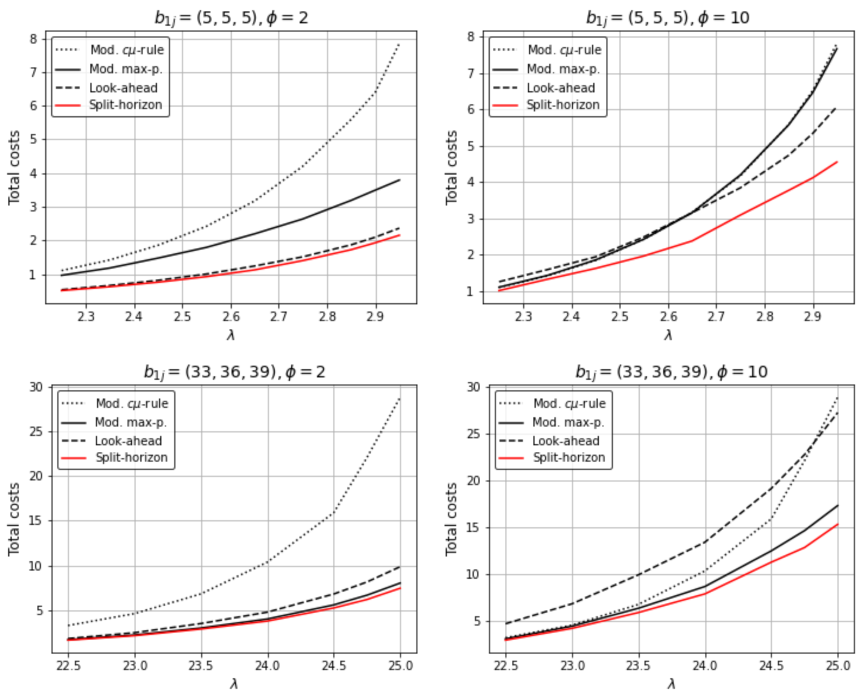


Figure 7.8: The split-horizon method compared to different routing policies on the SBR benchmark instances defined in Chen, Dong and Shi [44]. We have $I = 3, J = 3, N = 1, \psi_{ij} = 1, \mu_{ij} = 0.25, \lambda_i = c_i\bar{\lambda}_i, c_1 = 0.29, c_2 = 0.33, c_3 = 0.38, h_i = 1, f_{ij} = 0$ if $i == j, \phi$ else. For the split-horizon, we set $\Delta = 1$ and $T = 0$.

In Figure 7.8, the split-horizon method consistently outperforms all other policies across the entire range of $\bar{\lambda}$. No other allocation policy demonstrates such robust results. Specifically, the look-ahead policy excels in small-sized instances for most parameter values, while the modified maximum pressure policy performs best in large-sized instances. Thus, the split-horizon method effectively captures the key factors influencing overflow decisions, maintaining superior performance regardless of parameter settings.

7.5.2 Comparison to HSR benchmark

First, we compare the split-horizon method to the best-known policy for HSR by Zychlinski, Chan and Dong [202]. This study involves a single queue with customers having different service requirements, leading to $J = 1$ and $N = 1$. For the non-preemptive case, the researchers propose the idle-aware rule, an optimization program balancing server occupancy with class i customers for the highest instantaneous cost reduction and reducing idleness using a Γ term. However, Γ lacks a robust value due to its dependency on specific parameters. Therefore, three priority rules were defined for the two-server, two-class setting, i.e. $I = 2, \psi_{11} = 2, \psi_{21} = 1$, and $b_{11} = 2$. The priority rules are (1) P_1: prioritize class 1, (2) P_2: prioritize class 2, and (3) P_2^1: prioritize class 2 if two class 2 customers can be placed. We evaluate the split-horizon method against these priority rules using a benchmark instance from Zychlinski, Chan and Dong [202], as shown in Figure 7.9.

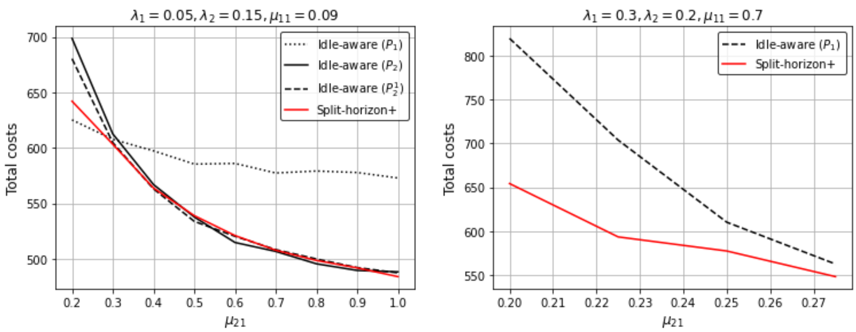


Figure 7.9: Split-horizon+ method compared to different idle-aware policies on the HSR benchmark instance defined in Zychlinski, Chan and Dong [202]. We have $I = 2, J = 1, N = 1, \boldsymbol{b}_1 = (2), \boldsymbol{\psi}_1 = (2), \boldsymbol{\psi}_2 = (1), \boldsymbol{\lambda} = (0.05, 0.15), \boldsymbol{f}_i = (0), \boldsymbol{h} = (5, 1)$. For the split-horizon+, we set $\Delta = 1, T = 9$, and $\xi = 0.1$.

In the left graph of Figure 7.9, it can be seen that for different μ_{21}, different priority settings perform well. We find that the split-horizon performs robustly for the complete range of μ_{21}-values. We see that the split-horizon is able to capture the priority switch based on the μ_{21}-value. For $\mu_{21} < 0.3$, the split-horizon prioritizes class 1. However, as the queue for class 2 becomes large, the split-horizon starts placing class 2 customers, which explains the small increase in costs value compared to always prioritizing class 1. Moreover, at $\mu_{21} = 0.3$, the split-horizon captures the P_2^1-rule, and for $\mu_{21} > 0.4$, the split-horizon prioritizes class 2.

In the priority settings based on the idle-aware rule, customers in the queue are always placed if feasible. However, there are instances where it is optimal to wait for a more important incoming customer, applying strategic idleness. In such cases, as shown in the right plot of Figure 7.9, the split-horizon method outperforms all priority settings. Only P_1 is displayed as it is the next best option. Additionally, when the system is busier (i.e., μ_{21} is low), strategic idleness becomes more critical, resulting in larger cost differences compared to scenarios where μ_{21} is higher.

To further illustrate the importance of strategic idleness, consider Figure 7.10, which shows a potential sample path. Despite class 1 being prioritized by P_1, resource constraints lead to an excessively long queue for class 1. This underscores the need for a scheduling policy with strategic idleness to achieve robust scheduling. The split-horizon method balances queue lengths across customer classes, optimizing waiting times.

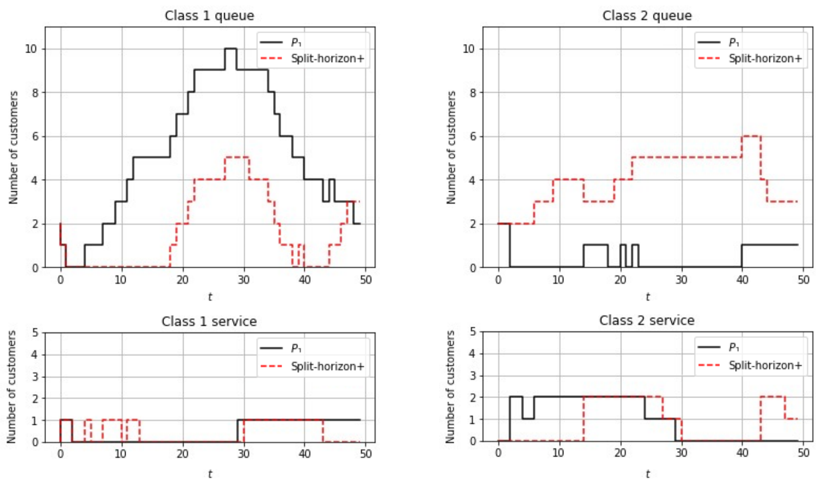


Figure 7.10: Sample paths of P_1 and the split-horizon+ method for the same trace for the HSR instance $I = 2, J = 1, N = 1, \boldsymbol{b}_1 = (2), \boldsymbol{\psi}_1 = (2), \boldsymbol{\psi}_2 = (1), \boldsymbol{\mu}_1 = (0.7), \boldsymbol{\mu}_2 = (0.2), \boldsymbol{\lambda} = (0.3, 0.2), \boldsymbol{f}_i = (0), \boldsymbol{h} = (5, 1)$. For the split-horizon+, we set $\Delta = 1, T = 9$, and $\xi = 0.1$.

7.5.3 Optimality gaps for the KBR benchmark

To evaluate the performance of the split-horizon method for a KBR model, we conducted an experiment with instances generated as follows: There are four classes ($I = 4$) and two server pools ($J = 2$). Classes 1 and 2 have fixed resource requirements $\psi_{11} = \psi_{21} = 1$, while classes 3 and 4 have resource requirements $\psi_{31} = \psi_{41} = \{2, 3, 4\}$. Classes 1 and 3 prefer server pool 1, and classes 2 and 4 prefer server pool 2, requiring simultaneous packing and routing decisions. We defined varying load settings $\rho_i = \{0.7, 0.8, 0.9\}$, implemented by calculating s_i^* via (7.12) and selecting $\lambda_i = \rho_i s_i^*$. This resulted in 27 small instances.

All scenarios were run in a Discrete Event Simulation (DES) using the batch mean method. After a warming-up period of 10^5 evaluated clients, 100 batches of 10^5 clients each were run, resulting in a maximum standard error of 0.005 for the mean long-term average costs for all runs. Four allocation policies were evaluated: (i) greedy, which equals the split-horizon method for only the finite horizon and $T = 0$, (ii) split-horizon, (iii) split-horizon+, and (iv) the adjusted look-ahead, which is the look-ahead method given in Chen, Dong and Shi [44] including resource constraints. The optimality gaps for the different policies are given in Figure 7.11.

Figure 7.11: Optimality gaps (left) and waiting times (right) for different allocation policies. The policies were evaluated on 27 KBR instances with fixed $I = 4, J = 2, N = 1, \boldsymbol{b}_j = (4), \boldsymbol{\psi}_1 = \boldsymbol{\psi}_2 = (1, 1), h_1 = h_2 = 1, \boldsymbol{\mu}_i = (1, 1), c_i = 0.25, \boldsymbol{f}_1 = \boldsymbol{f}_3 = (0, 2), \boldsymbol{f}_2 = \boldsymbol{f}_4 = (2, 0)$. Other parameters vary: $\rho_i = \{0.7, 0.8, 0.9\} \forall i \in I, \boldsymbol{\psi}_3 \in \{(2, 2), (3, 3), (4, 4)\}, \boldsymbol{\psi}_4 = \boldsymbol{\psi}_3, h_3 \in \{2, 3, 4\}, h_4 = h_3$. We set $Q_{max} = 5$ and $c_{\text{rej}} = 2$. Finally, $\Delta = \frac{1}{\sum_{i=1}^{I}(\lambda_i + \sum_{j=1}^{J} b_{j1}\mu_{ij})}, T = \lceil \frac{1.5}{\min_i \lambda_i \Delta} \rceil$, and $\xi = 0.1$.

We find that the split-horizon and split-horizon+ methods exhibit optimality gaps of 2.5% and 2.3%, respectively, thereby outperforming the greedy policy, which has an optimality gap of 32.6%, and the adjusted look-ahead policy, which has an optimality gap of 10.4%. The poor performance of the greedy policy highlights the importance of considering future events. Additionally, the split-horizon+ method shows a slight improvement over the split-horizon method, indicating a marginal benefit of accounting for future arrivals. We also evaluated the aggregated versions of the split-horizon and split-horizon+ methods, which resulted in optimality gaps of 2.6% and 3.4%, respectively, demonstrating strong performance as well. For further insight into the optimal placement policy compared to the split-horizon+ method, we provide several policy tables for various states in Appendix 7.A.2.

Furthermore, we compare the class-specific waiting times between the split-horizon+ and adjusted look-ahead policies. The results, provided in the right graph in Figure 7.11, demonstrate that the split-horizon+ method not only minimizes long-term costs more effectively but also achieves more balanced waiting times across different classes. Given that balanced waiting times are a crucial criterion for evaluating the efficacy of a policy in the mental health context, these findings underscore the superior performance of our method in this regard.

7.5.4 Case study

For our case study, we identified all group homes in Amsterdam, comprising 32 homes with a total of 143 beds. These homes are distributed across ten locations, as shown in Figure 7.12. Additionally, we accessed health data from Statistics Netherlands on individuals with severe mental illness living in community care in Amsterdam [39]. The average age in this type of housing is 49.8 years, with 35% identifying as female. Prior to their admission, 40.4% received therapy for schizophrenia, 21.3% for addiction, 8.5% for mood disorders, and 6.4% for personality disorders. We found no seasonal trends in the arrival patterns, allowing us to assume Poisson arrivals. Analysis of service duration data indicates an average length-of-stay of approximately ten years. Due to privacy constraints, we could not obtain specific client characteristics, necessitating assumptions about the potential patient population. Furthermore, five healthcare organizations provide these facilities, for which we assume that centralizing the waiting list is feasible. Therefore, we selected a fraction of the homes for the case study.

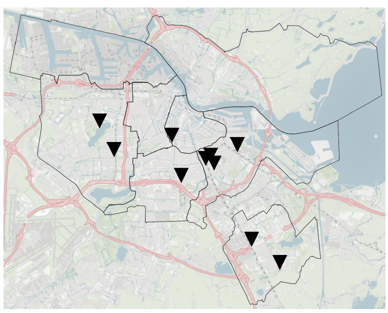


Figure 7.12: Locations of group homes in Amsterdam.

The current placement procedure is as follows. A potential client, assisted by an employee of a care organization or the municipality, evaluates the available homes. During this process, the client's restrictions and preferences are thoroughly discussed. Together, they identify a few best-suited homes, typically two, and the client is placed on the waiting list of the home with the shortest queue. This approach can be viewed as a d-policy with $d = 2$ [121]. When a bed becomes available, the first person in the queue who meets the group compatibility requirements is admitted, following a 'First-Come-First-Feasible-Served' policy. We denote this as the Current Policy (CP).

Our proposed alternative to CP is the split-horizon method, where we select the two best-suited homes as preferred options. If a client is placed in a home other than these preferred options, they incur an off-placement penalty ϕ. As we have five houses, for the overflow costs we denote $\sigma_i(\phi, \phi, \phi, 0, 0)$ as the i^{th} permutation consisting of three overflow penalties ϕ and two preferred values 0. Given the multiple preferred house choices, we employ the version of the IQP with class aggregation. The comparison between the CP and the split-horizon method is provided in Figure 7.13.

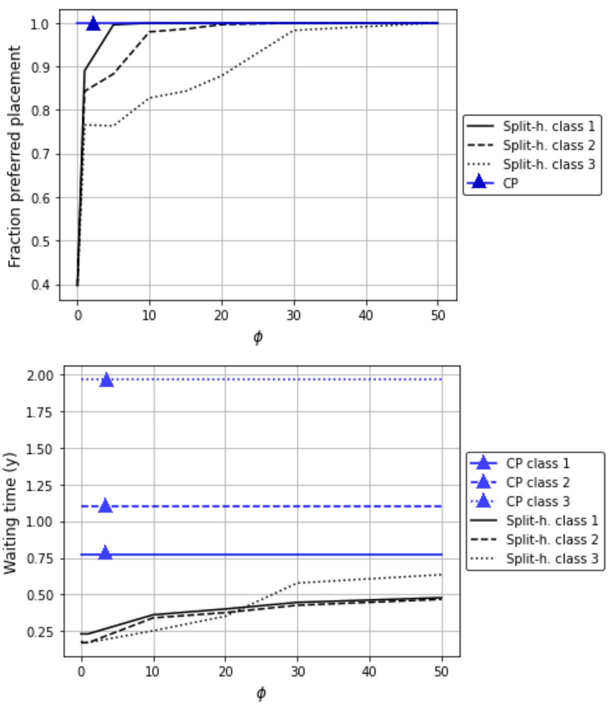


Figure 7.13: Results of the case study. The instance is $I = 30, J = 5, N = 2, \psi_{i1} = 1, \psi_{i2} = 1$ and $\boldsymbol{f}_i = \sigma_i$ if $i \in \{1, ..., 10\}, \psi_{i2} = 2$ and $\boldsymbol{f}_i = \sigma_{i-10}$ if $i \in \{11, ..., 20\}, \psi_{i2} = 3$ and $\boldsymbol{f}_i = \sigma_{i-20}$ if $i \in \{21, ..., 30\}, \boldsymbol{b}_j = (5, 8), \mu_{ij} = 0.1, \boldsymbol{h} = (1, 2, 3), \boldsymbol{\lambda} = (1.4, 0.4, 0.2)$. We set $\Delta = 0.22$ and $T = 10$.

In Figure 7.13, the heuristic places a higher fraction of individuals in their preferred houses as opportunity costs increase. Classes with lower resource requirements are placed more frequently in preferred houses, ensuring all individuals are accommodated in a preferred house when opportunity costs $\phi \geq 50$. However, this prioritization leads to longer waiting times, as shown in the bottom figure, which worsens for the split-horizon method as opportunity costs rise. Furthermore, waiting times under the current policy can be excessively long, with hard-to-place individuals experiencing delays of up to two years, consistent with estimates from mental healthcare organizations.

In the same figure, we observe a significant imbalance in waiting times among the three classes under the current policy, where higher resource requirements result in longer waits. It is evident that hard-to-place clients are prioritized less effectively compared to easy-to-place clients when no scheduling policy is used. In contrast, the split-horizon method employs an advanced scheduling policy that significantly improves the balance of waiting times across classes, achieving much shorter waits than the current policy. Thus, these findings underscore the inefficiency and inequity of the current policy, highlighting the potential enhancements offered by the split-horizon method.

7.6 Discussion

In this chapter, we address the problem of allocating individuals to mental healthcare residential facilities. We formulate this problem as a KBR model, where customer classes are routed to server pools with resource constraints and preferences. Designing an effective policy for this problem is complex and requires a scheduling approach, as we show that index-based priority settings can lead to opposite outcomes due to varying resource demands. We propose the split-horizon method, which divides the look-ahead horizon into short-term and long-term planning. The goal is to balance favorable scheduling decisions in the short-term and the effects of these decisions on the long-term. We find that the split-horizon method performs well, with an optimality gap of 2.3%, and outperforms state-of-the-art policies on benchmark instances for skill-based routing with overflow and queues with heterogeneous service requirements.

We applied our split-horizon method to a real-life case study involving the allocation of individuals with severe mental illness to group homes in Amsterdam, the Netherlands. We found that the method demonstrated superior performance over the existing placement policy by achieving shorter waiting times. Moreover, due to its threshold-based actions, the split-horizon method yields more balanced waiting times between easy-to-place and hard-to-place customer classes. With our methodology, we enhance the operations research literature in the mental health domain by incorporating domain-specific elements such as clustering customers based on resource utilization and prioritizing acceptable waiting times [129]. Given the significant logistical challenges in mental healthcare, we advocate for increased application of operations research methodologies in this highly pertinent, yet underexplored, domain.

We suggest several directions for future research. First, gaining theoretical insights into the model would be valuable. For instance, can we explicitly determine which customer class has priority? In what specific cases is it optimal to apply strategic idleness or to overflow? Currently, these insights are lacking due to the complexity of the problem, but obtaining them would pave the way for more efficient policies. Second, the model could be extended by incorporating elements such as abandonments or customers changing resource requirements during waiting or service. Finally, exploring other solution approaches, such as applying reinforcement learning to the KBR model, would be interesting and challenging due to the nonpreemptive nature and large state and action spaces. In conclusion, we consider the KBR model to be a general model with practical utility in various healthcare settings and call centers, making it worthwhile to investigate further.

7.A Appendix

7.A.1 MDP formulation

We now formulate the MDP for the KBR model. Recall that the state s is defined in (7.1). Let $\boldsymbol{q} = (q_1, \dots, q_I)$ be a vector of queue lengths and $\boldsymbol{p}$ be the matrix with entries p_{kj} denoting the number of class k customers in server pool j. Using (7.1), the state s can be described as $s = \{\boldsymbol{q}, \boldsymbol{p}\}$. For numerical purposes, we truncate the state space and set $Q_{\max}$ as the maximum queue length, giving the state space

$$\mathscr{S} = \Bigg\{ s \in \mathbb{N}^I \times \mathbb{N}^{K \times J} \mid \sum_{i=1}^{I} q_i \leq Q_{\max}, \sum_{k=1}^{K} \phi_{kn} p_{kj} \leq b_{jn}, \\ j = 1, \dots, J,\ n = 1, \dots, N \Bigg\}.$$

The set of feasible actions for state s is given by $A(s)$, as defined in (7.2). Let N_j be the maximum number of customers at server pool j, which is assumed to be finite. As the first resource is associated with the number of servers, we typically have $\phi_{i1} = 1$, for all $i = 1, \dots, I$, in which case $N_j \leq b_{1j} < \infty$. Without loss of generality, we rescale time such that $\sum_{i=1}^{I} \lambda_i + \sum_{j=1}^{J} N_j \max\{\mu_{ij}\} = 1$.

To define the transition probabilities, let s be the preaction state and $s^* = \{\boldsymbol{q}^*, \boldsymbol{p}^*\}$ the postaction state. In particular, using action $a \in A(s)$ in state s, the postaction state s^* has entries $q_i^* = q_i - \sum_j a_{ij}$ and $p_{kj}^* = p_{kj} + \sum_i \gamma_{ik} a_{ij}$. Let e_i denote the i-th unit vector and e_{kj} be a matrix with zeroes and a 1-entry in the i-th row and the j-th column. The transition probabilities to state $s' = \{\boldsymbol{q}', \boldsymbol{p}'\}$, using the postaction states $s^* = \{\boldsymbol{q}^*, \boldsymbol{p}^*\}$, are given by

$$P(s' \mid s,a) = \begin{cases} \sum_{i=1}^{I} \lambda_i & \text{if } s' = s^*, \sum_{i=1}^{I} q_i^* = Q_{\max}, \\ \lambda_i & \text{if } \boldsymbol{q}' = \boldsymbol{q}^* + e_i, \sum_{i=1}^{I} q_i^* < Q_{\max}, \\ & \quad \boldsymbol{p}' = \boldsymbol{p}^*, \\ \mu_{kj} p_{kj}^* & \text{if } \boldsymbol{q}' = \boldsymbol{q}^*, \boldsymbol{p}' = \boldsymbol{p}^* - e_{kj}, \\ 1 - \sum_{k=1}^{K} \sum_{j=1}^{J} \mu_{kj} p_{kj}^* - \sum_{i=1}^{I} \lambda_i & \text{if } s' = s^*, \\ 0 & \text{otherwise.} \end{cases}$$

The one-epoch cost associated with action a and preaction state s are

$$c(s,a) = \sum_{i=1}^{I} \sum_{j=1}^{J} f_{ij} a_{ij} + \sum_{i=1}^{I} h_i \left(q_i - \sum_{j=1}^{J} a_{ij} \right) + c_{\text{rej}} \bar{\lambda} \mathbb{1}_{\{\sum_i q_i - \sum_{i,j} a_{ij} = Q_{\max}\}},$$

with c_{rej} the cost for rejecting a customer due to the upper bound on the queue length. Note that the holding and rejection costs apply to the postaction state.

For the long-run average cost, we need to solve the Bellman equations

$$g^* + V^*(s) = \min_{a \in A(s)} \left\{ c(s,a) + \sum_{s' \in \mathscr{S}} P(s' \mid s,a) V^*(s') \right\}, \qquad s \in \mathscr{S}.$$

Here, g^* is the optimal long-run average cost and $V^*(\cdot)$ is the optimal relative value function. For small instances, we numerically solve the Bellman equations using value iteration. Clearly, given the $(I + KJ)$-dimensional state space, solving such an MDP is computationally prohibitive for any reasonably sized instance.

7.A.2 Policy tables for KBR benchmark instance

To gain deeper insights into the policy structure of the split-horizon+ method, we analyze the policy tables comparing the optimal policy with the split-horizon+ method, as shown in Figure 7.14. We focus on a specific instance from the KBR benchmark set, featuring four customer classes ($I = 4$) allocated to two server pools ($J = 2$). Initially, we select three 'base' states where server pool 1 is partially filled while server pool 2 remains empty. We vary the queue lengths for customer classes 1 and 3, which are dedicated to server pool 1 and potentially overflow into the vacant server pool 2.

For states where server pool 1 is empty, specifically $s = (q_1, 0, q_3, 0,\ 0, 0, 0, 0)$, the split-horizon+ method accurately determines the optimal policy for q_1 and q_3, where $q_1 + q_3 \leq 5$. Notably, the method prioritizes customer class 3 appropriately in server pool 1, maintaining strict precedence in scenarios without overflow. Additionally, an overflow threshold is identified at $q_3 = 3$, beyond which overflow becomes necessary if $q_1 \geq 3$. The split-horizon+ method precisely identifies this threshold and the optimal placement and overflow strategy, exemplified in state $s = (3, 0, 1, 0, 0, 0, 0, 0)$, where the action matches the optimal action $a = (2, 0, 1, 0, 0, 0, 1, 0)$.

For states $s = (q_1, 0, q_3, 0, 2, 0, 0, 0)$ and $s = (q_1, 0, q_3, 0, 4, 0, 0, 0)$, representing busier system conditions, overflow decisions occur more frequently. Here, the split-horizon+ method consistently makes accurate placement and overflow decisions. The policy tables in Figure 7.14 highlight instances (colored cells) where the split-horizon method exhibits a slightly more conservative approach, overshooting the optimal overflow threshold. For instance, in states $s = (0, 0, 1, 0, 4, 0, 0, 0)$, overflow should have commenced, whereas the split-horizon+ method initiates overflow at $s = (0, 0, 2, 0, 4, 0, 0, 0)$ and $s = (1, 0, 1, 0, 4, 0, 0, 0)$.

In conclusion, the split-horizon+ method performs accurately in determining customer prioritization and overflow decisions. It avoids strategic idleness by maximizing utilization of dedicated server pool 1, aligning well with the optimal policy for this parameter configuration.

Optimal policy

$s = (q_1, 0, q_3, 0, 0, 0, 0, 0)$

q_3 \ q_1	0	1	2	3	4	5
2	(0,0,2,0,0,0,0,0)	(0,0,2,0,0,0,0,0)	(0,0,2,0,0,0,0,0)	(2,0,1,0,0,0,1,0)		
1	(0,0,1,0,0,0,0,0)	(1,0,1,0,0,0,0,0)	(2,0,1,0,0,0,0,0)	(2,0,1,0,0,0,0,0)	(2,0,1,0,0,0,0,0)	
0		(1,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)	(3,0,0,0,0,0,0,0)	(4,0,0,0,0,0,0,0)	(4,0,0,0,0,0,0,0)

$s = (q_1, 0, q_3, 0, 2, 0, 0, 0)$

q_3 \ q_1	0	1	2	3	4	5
2	(0,0,1,0,0,0,1,0)	(0,0,1,0,0,0,1,0)	(0,0,1,0,0,0,1,0)	(2,0,0,0,0,0,2,0)		
1	(0,0,1,0,0,0,0,0)	(0,0,1,0,0,0,0,0)	(0,0,1,0,0,0,0,0)	(2,0,0,0,0,0,1,0)	(2,0,0,0,0,0,1,0)	
0		(1,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)

$s = (q_1, 0, q_3, 0, 4, 0, 0, 0)$

q_3 \ q_1	0	1	2	3	4	5
2	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,2,0)	(0,0,0,0,0,0,2,0)	(0,0,0,0,0,0,2,0)		
1	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,1,0)	
0		(0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0)	(0,0,0,0,1,0,0,0)

Split-horizon+

$s = (q_1, 0, q_3, 0, 0, 0, 0, 0)$

q_3 \ q_1	0	1	2	3	4	5
2	(0,0,2,0,0,0,0,0)	(0,0,2,0,0,0,0,0)	(0,0,2,0,0,0,0,0)	(2,0,1,0,0,0,1,0)		
1	(0,0,1,0,0,0,0,0)	(1,0,1,0,0,0,0,0)	(2,0,1,0,0,0,0,0)	(2,0,1,0,0,0,0,0)	(2,0,1,0,0,0,0,0)	
0		(1,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)	(3,0,0,0,0,0,0,0)	(4,0,0,0,0,0,0,0)	(4,0,0,0,0,0,0,0)

$s = (q_1, 0, q_3, 0, 2, 0, 0, 0)$

q_3 \ q_1	0	1	2	3	4	5
2	(0,0,1,0,0,0,0,0)	(0,0,1,0,0,0,1,0)	(0,0,1,0,0,0,1,0)	(2,0,0,0,0,0,2,0)		
1	(0,0,1,0,0,0,0,0)	(0,0,1,0,0,0,0,0)	(0,0,1,0,0,0,0,0)	(2,0,0,0,0,0,1,0)	(2,0,0,0,0,0,1,0)	
0		(1,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)	(2,0,0,0,0,0,0,0)

$s = (q_1, 0, q_3, 0, 4, 0, 0, 0)$

q_3 \ q_1	0	1	2	3	4	5
2	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,2,0)	(0,0,0,0,0,0,2,0)	(0,0,0,0,0,0,2,0)		
1	(0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,1,0)	(0,0,0,0,0,0,1,0)	
0		(0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0)

Figure 7.14: Action tables for $I = 4, K = 2, J = 2, N = 1, \boldsymbol{b}_j = 4, \boldsymbol{\psi}_1 = \boldsymbol{\psi}_2 = (1,1), \boldsymbol{\psi}_3 = \boldsymbol{\psi}_4 = (2,2), \boldsymbol{\mu}_i = (1,1), \rho_i = 0.7, \boldsymbol{h} = (1,1,2,2), \boldsymbol{f}_1 = \boldsymbol{f}_3 = (0,2), \boldsymbol{f}_2 = \boldsymbol{f}_4 = (2,0)$. For other parameters, see Figure 7.11. States are denoted by $(q_1, q_2, q_3, q_4, p_{11}, p_{21}, p_{12}, p_{22})$ and actions by $(a_{11}, a_{21}, a_{31}, a_{41}, a_{12}, a_{22}, a_{32}, a_{42})$. Actions in which the heuristic differs from the optimal policy are colored.

8 Implications and Outlook

This thesis is based on an interdisciplinary collaboration between two heretofore separate domains: mathematical modelling and stochastic optimization on the one hand, and the elderly care domain on the other hand [191]. With a focus on patient flows, capacity, and efficient allocation, various critical aspects of the current healthcare system have been identified that require an interdisciplinary approach. As the aging population puts increasing pressure on the healthcare system, the need for a logistical perspective will become even more crucial in the next decades.

8.1 Policy Implications

This thesis focuses on two key areas: short-term care and long-term care. An investigation of both components of the healthcare system for old age yielded possible directions for policy:

- *Ensure 24/7 admission capabilities in Intermediate Care*: It is essential to improve round-the-clock admission options in regions where these services are currently insufficient. This will reduce unnecessary hospitalizations, which are costly and may lead to adverse events.
- *Develop a regional approach to Long-Term Care waiting lists*: A regional approach to managing waiting lists in nursing homes (or home care) should be developed, so that waiting times can be reduced and efficient allocation can be leveraged.

For short-term care, particularly Intermediate Care (IC), this thesis explored strategies to reduce waiting times and minimize unnecessary hospital admissions. A simulation study conducted in the Amsterdam region revealed that the waiting times were *not* due to a lack of beds but rather due to insufficient admission capacity during evenings and weekends [12]. A survey study showed that some IC facilities have only a few beds in nursing homes, complicating acute admissions [171]. For regions where this is still inadequately addressed, we recommend investing in a more efficient 24/7 admission process. This approach ensures that older adults can be cared for outside of the hospital, receiving the right care at the right place.

For long-term care, we advocate for a regional registration system for the waiting lists. At some regions, successful pilot programs have already implemented for this issue [5]. A regional system can create the necessary economies of scale to reduce waiting times. Additionally, we recommend adopting an allocation model to optimize the matching of older individuals to available spaces, taking into account their preferences and flexibility. A points-based system, similar to that used in social housing allocation, could also be beneficial for this purpose.

8.2 Potential for a Macro Model

Our research examined how reducing waiting times across various care providers, such as nursing homes and IC, could be achieved. Reducing these waiting times not only impacts the providers themselves but also affects system-wide capacity, as bed-blocking in upstream providers occurs when transitions are delayed. By maintaining optimal capacity at the next care level, patient flow throughout the system can be maximized.

To determine this 'optimal' capacity, it is necessary to understand how older adults move through the healthcare landscape, considering both transition probabilities and Length-of-Stay (LoS). In Chapter 2, we mapped the patient journeys of older persons, revealing that those aged 85+ experience multiple care transitions within a single pathway. A model based on factors like age or frailty status [188] would be useful in this context. We also developed a method to estimate the distribution of LoS, with a focus on longer durations using uncensored data [176]. Together, these elements could form a model for determining the optimal (regional) capacity for each type of care.

How this 'macro model' should be structured is a question for future research. Initial attempts, such as system dynamics [58] and Jackson networks [176], were too high-level to capture the necessary and complex dynamics. An alternative approach could be to combine the models developed in this thesis for short-term care [12] and long-term care [11] into a single framework. The short-term care model focuses on urgency and timeframes of hours/days, while the long-term care model emphasizes preferences and timeframes of months/years. Future work should explore if and how these models can be integrated.

8.3 Generalization to Other Care Domains

The models developed in this thesis are generic in nature and can be applied to other healthcare domains beyond care for old age. A good example is the allocation model developed for residential psychiatric care in this thesis, which emerged as an extension of the nursing home allocation model. The new constraints added in this extension could also be directly applied to healthcare for older adults, for instance, in the context of multimorbidity. Although the model is general, the data and parameter values will ultimately make sure that the model is tailored to the application setting.

Potential applications in other care sectors include youth care, maternity care, psychiatric care, and disability care. The allocation model for psychiatric residential care could be adapted to residential facilities in disability care. Many of these sectors currently face long waiting lists. In psychiatric care, waiting lists have become so extensive that patients are now taking legal action against the State and insurers [131]. Prolonged waiting times often result in the deterioration of conditions, reducing the likelihood of timely recovery and reintegration into society, such as returning to work or securing employment [186]. Hence, insights into optimizing capacity utilization in this domain are crucial as well.

8.4 Final Thoughts

Several recommendations emerged indirectly from this thesis. One key suggestion is to improve the registration of *desired* care pathways at the national level, rather than just the *realized* pathways. It is possible that demand for certain types of care is higher than recorded, but due to capacity limitations, faster alternatives are often chosen instead. Additionally, in hospitals, incorrect bed days are currently registered, which is helpful to understand the issue of bed-blocking. Similar initiatives could be useful in other care settings.

Finally, several promising initiatives are currently underway in elderly care that align with logistical efficiency, such as the development of Integral Capacity Management departments in hospitals and the emerging Care Coordination Centers (ZCCs in Dutch). These centers, much like the National Coordination Center for Patient Distribution during the COVID-19 pandemic, can make decisions about patient distribution at the regional level. We hope that ZCCs will continue to play a similar role, with a focus on the efficient assignment of older adults to care locations. Furthermore, attention should be given to the organization of care forms themselves. For example, IC is organized differently across the country, sometimes as a separate ward and other times as a single bed within a nursing home [171]. Care models that are too small in scale may hinder an efficient admission process and limit opportunities for economies of scale, indicating significant potential for improvement in logistics.

List of Publications

Arntzen, R. J., Bekker, R., & van der Mei, R. D. (2024a). Knapsack-based routing for mental health placements: A split-horizon approach. *In review*.

Arntzen, R. J., Bekker, R., & van der Mei, R. D. (2024b). Preference-based allocation of patients to nursing homes. *Operations Research for Health Care*, *42*, 100442.

Arntzen, R. J., Bekker, R., Smeekes, O. S., Buurman, B. M., Willems, H. C., Bhulai, S., & van der Mei, R. D. (2022). Reduced waiting times by preference-based allocation of patients to nursing homes. *Journal of the American Medical Directors Association*, *23*(12), 2010–2014.

Arntzen, R. J., van den Besselaar, J. H., Bekker, R., Buurman, B. M., & van der Mei, R. D. (2023). Avoiding hospital admissions and delayed transfers of care by improved access to intermediate care: A simulation study. *Journal of the American Medical Directors Association*, *24*(7), 945–950.

van den Besselaar, J. H., Arntzen, R. J., MacNeil Vroomen, J. L., Hertogh, C. M. P. M., & Buurman, B. M. (2023). Short term residential care in the Netherlands: Patient and facility characteristics from a national database and survey. *Submitted*.

de Boer, T. R., Arntzen, R. J., Bekker, R., Buurman, B. M., Willems, H. C., & van der Mei, R. D. (2025). Process mining on national healthcare data for the discovery of patient journeys of older adults. *Journal of the American Medical Directors Association*, *26*(1), 105333.

Bibliography

[1] Abrahamsen, J. F., Haugland, C., Nilsen, R., & Ranhoff, A. (2016). Three different outcomes in older community-dwelling patients receiving intermediate care in nursing home after acute hospitalization. *The Journal of Nutrition, Health and Aging*, *20*(4), 446–452.

[2] ActiZ. (2022). *Hoe staat het met de wachtlijsten voor verpleeghuizen? (in Dutch)*. Retrieved August 30, 2024, from https://www.actiz.nl/sites/default/files/2022-08/Infographic%20Hoe%20staat%20het%20met%20de%20wachtlijsten%20voor%20verpleeghuizen%20juni%202022.pdf

[3] ActiZ. (2024). Infographic ELV (in Dutch).

[4] *Actiz (in Dutch)*. (no date). Retrieved August 30, 2024, from https://www.actiz.nl/

[5] Algemeen Dagblad. (2022). *Verpleeghuizen in regio werken met één wachtlijst, zodat kwetsbare ouderen sneller een plek krijgen (in Dutch)*. Retrieved August 30, 2024, from https://www.ad.nl/dordrecht/verpleeghuizen-in-regio-werken-met-een-wachtlijst-zodat-kwetsbare-ouderen-sneller-een-plek-krijgen~a5d3962f/

[6] Altman, E., Jiménez, T., & Koole, G. (2001). On optimal call admission control in resource-sharing system. *IEEE Transactions on Communications*, *49*(9), 1659–1668.

[7] Andersen, A. R., Nielsen, B. F., & Reinhardt, L. B. (2017). Optimization of hospital ward resources with patient relocation using Markov chain modeling. *European Journal of Operational Research*, *260*(3), 1152–1163.

[8] Ariss, S. M., Enderby, P. M., Smith, T., Nancarrow, S. A., Bradburn, M. J., Harrop, D., Parker, S. G., McDonnell, A., Dixon, S., Ryan, T., Haymann, A., & Campbell, M. (2015). Secondary analysis and literature review of community rehabilitation and intermediate care: An information resource. *Health Services and Delivery Research*, *3.1*.

[9] Arntzen, R. J., Bekker, R., & van der Mei, R. D. (2024a). Knapsack-based routing for mental health placements: A split-horizon approach. *In review*.

[10] Arntzen, R. J., Bekker, R., Smeekes, O. S., Buurman, B. M., Willems, H. C., Bhulai, S., & van der Mei, R. D. (2022). Reduced waiting times by preference-based allocation of patients to nursing homes. *Journal of the American Medical Directors Association*, *23*(12), 2010–2014.

[11] Arntzen, R. J., Bekker, R., & van der Mei, R. D. (2024b). Preference-based allocation of patients to nursing homes. *Operations Research for Health Care*, *42*, 100442.

[12] Arntzen, R. J., van den Besselaar, J. H., Bekker, R., Buurman, B. M., & van der Mei, R. D. (2023). Avoiding hospital admissions and delayed transfers of care by improved access to intermediate care: A simulation study. *Journal of the American Medical Directors Association*, *24*(7), 945–950.

[13] Atar, R., Giat, C., & Shimkin, N. (2010). The cμ/θ-rule for many-server queues with abandonment. *Operations Research*, *58*(5), 1427–1439.

[14] Atella, V., Piano Mortari, A., Kopinska, J., Belotti, F., Lapi, F., Cricelli, C., & Fontana, L. (2019). Trends in age-related disease burden and healthcare utilization. *Aging Cell*, *18*(1), e12861.

[15] Australian Government. (2020). Report on government services: Aged care services.

[16] Baker, G. R., Norton, P. G., Flintoft, V., Blais, R., Brown, A., Cox, J., et al. (2004). The Canadian Adverse Events Study: The incidence of adverse events among hospital patients in Canada. *Canadian Medical Association Journal*, *170*(11), 1678–1686.

[17] Bakx, P., Wouterse, B., Van Doorslaer, E., & Wong, A. (2020). Better off at home? Effects of nursing home eligibility on costs, hospitalizations and survival. *Journal of Health Economics*, *73*, 102354.

[18] Bekker, R., Moeke, D., & Schmidt, B. (2019). Keeping pace with the ebbs and flows in daily nursing home operations. *Health Care Management Science*, *22*(2), 350–363.

[19] Bell, S. L., & Williams, R. J. (2001). Dynamic scheduling of a system with two parallel servers in heavy traffic with resource pooling: Asymptotic optimality of a threshold policy. *The Annals of Applied Probability*, *11*(3), 608–649.

[20] Ben-Tovim, D. I., Dougherty, M. L., O'Connell, T. J., & McGrath, K. M. (2008). Patient journeys: The process of clinical redesign. *Medical Journal of Australia*, *188*(S6), S14–S17.

[21] Beresford, B. A., Mann, R. C., Parker, G. M., Kanaan, M., Neves De Faria, R. I., Rabiee, P., Weatherly, H. L. A., Clarke, S. E., Mayhew, E. T., Duarte, A. I., et al. (2019). Reablement services for people at risk of needing social care: The more mixed-methods evaluation. *Health Services and Delivery Research*, 1–254.

[22] Bergeron, P., Courteau, J., & Vanasse, A. (2015). Proximity and emergency department use: Multilevel analysis using administrative data from patients with cardiovascular risk factors. *Canadian Family Physician*, *61*(8), e391–e397.

[23] Bertazzoni, G., Cristofani, M., Ponzanetti, A., Trabalzini, A., Attalla, H., De Vito, C., & Villari, P. (2008). Scant justification for interhospital transfers: A cause of reduced efficiency in the emergency department. *Emergency Medicine Journal*, *25*(9), 558–561.

[24] Bertsekas, D. (2021). *Rollout, policy iteration, and distributed reinforcement learning*. Athena Scientific.

[25] Bigby, C., & Beadle-Brown, J. (2018). Improving quality of life outcomes in supported accommodation for people with intellectual disability: What makes a difference? *Journal of Applied Research in Intellectual Disabilities*, *31*(2), e182–e200.

[26] Bigelow, D. A., Cutler, D. L., Moore, L. J., McComb, P., & Leung, P. (1988). Characteristics of state hospital patients who are hard to place. *Psychiatric Services*, *39*(2), 181–185.

[27] Blake, J. T. (2005). Shooting arrows in the dark: The policies and practices of waitlist management in Canada. *Clinical and Investigative Medicine*, *28*(6), 308–311.

[28] Bo, Y., Dawande, M., Huh, W. T., Janakiraman, G., & Nagarajan, M. (2019). Determining process capacity: Intractability and efficient special cases. *Manufacturing & Service Operations Management*, *21*(1), 139–153.

[29] Bowers, J. A. (2011). Simulating waiting list management. *Health Care Management Science*, *14*(3), 292–298.

[30] Boyd, C. M., Landefeld, C. S., Counsell, S. R., Palmer, R. M., Fortinsky, R. H., Kresevic, D., Burant, C., & Covinsky, K. E. (2008). Recovery of activities of daily living in older adults after hospitalization for acute medical illness. *Journal of the American Geriatrics Society*, *56*(12), 2171–2179.

[31] Brolin, R., Rask, M., Syrén, S., Baigi, A., & Brunt, D. A. (2015). Satisfaction with housing and housing support for people with psychiatric disabilities. *Issues in Mental Health Nursing*, *36*(1), 21–28.

[32] Brunt, D., & Hansson, L. (2002). Characteristics of the social environment of small group homes for individuals with severe mental illness. *Nordic Journal of Psychiatry*, *56*(1), 39–46.

[33] *bupaR docs*. (no date). Retrieved October 19, 2023, from https://bupaverse.github.io/docs/index.html

[34] Burkell, J., Wright, A., Hoffmaster, B., & Webb, K. (1996). A decision-making aid for long-term care waiting list policies: Modelling first-come, first-served vs. needs-based criteria. *Healthcare Management Forum*, *9*(1), 35–39.

[35] Buurman, B. M., Hoogerduijn, J. G., de Haan, R. J., Abu-Hanna, A., Lagaay, A. M., Verhaar, H. J., Schuurmans, M. J., Levi, M., & de Rooij, S. E. (2011). Geriatric conditions in acutely hospitalized older patients: Prevalence and one-year survival and functional decline. *PloS one*, *6*(11), e26951.

[36] Byran, S., Glasby, J., Jagger, C., Kaambwa, B., Martin, G., Nancarrow, S., Parker, H., Parker, S., Regen, E., Wilson, A., et al. (2006). *A national evaluation of the costs and outcomes of intermediate care for older people*. University of Birmingham Health Services Management Team.

[37] Caley, M., & Sidhu, K. (2011). Estimating the future healthcare costs of an aging population in the UK: Expansion of morbidity and the need for preventative care. *Journal of Public Health*, *33*(1), 117–122.

[38] Carpenter, G. I., Kotiadis, K., & Mackenzie, M. (2003). An evaluation of intermediate care services for older people: Summary report.

[39] Centraal Bureau voor de Statistiek. (2023). *Maatwerk en microdata (in Dutch)*. Retrieved October 13, 2023, from https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata

[40] Centraal Bureau voor de Statistiek. (2024). *Ouderen (in Dutch)*. Retrieved August 30, 2024, from https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/leeftijd/ouderen

[41] *Central Committee on Research Involving Human Subjects*. (no date). Retrieved August 30, 2024, from https://english.ccmo.nl/

[42] Chafe, R., Coyte, P., & Sears, N. A. (2010). Improving the management of waiting lists for long term care. *Healthcare Management Forum*, *23*(2), 58–62.

[43] Chan, C. W., Dong, J., & Green, L. V. (2017). Queues with time-varying arrivals and inspections with applications to hospital discharge policies. *Operations Research*, *65*(2), 469–495.

[44] Chen, J., Dong, J., & Shi, P. (2020). A survey on skill-based routing with applications to service operations management. *Queueing Systems*, *96*, 53–82.

[45] Chen, J., Dong, J., & Shi, P. (2023). Optimal routing under demand surges: The value of future arrival rates. *Operations Research*.

[46] Close, J. C., Lord, S. R., Antonova, E. J., Martin, M., Lensberg, B., Taylor, M., Hallen, J., & Kelly, A. (2012). Older people presenting to the emergency department after a fall: A population with substantial recurrent healthcare use. *Emergency Medicine Journal*, *29*(9), 742–747.

[47] Committee on the Future Health Care Workforce for Older Americans. (2008). *Retooling for an aging America: Building the health care workforce*. National Academies Press.

[48] Connelly, L. G., & Bair, A. E. (2004). Discrete event simulation of emergency department activity: A platform for system-level operations research. *Academic Emergency Medicine*, *11*(11), 1177–1185.

[49] Covinsky, K. E., Pierluissi, E., & Johnston, C. B. (2011). Hospitalization-associated disability: “She was probably able to ambulate, but I’m not sure”. *Journal of the American Medical Association*, *306*(16), 1782–1793.

[50] Cox, D., & Smith, W. (1961). *Queues*. CRC Press.

[51] Cristoforetti, A., Gennai, F., & Rodeschini, G. (2011). Home sweet home: The emotional construction of places. *Journal of Aging Studies*, *25*(3), 225–232.

[52] Crotty, M., Whitehead, C. H., Wundke, R., Giles, L. C., Ben-Tovim, D., & Phillips, P. A. (2005). Transitional care facility for elderly people in hospital awaiting a long term care bed: Randomised controlled trial. *British Medical Journal*, *331*(7525), 1110–1114.

[53] Dai, J. G., & Lin, W. (2005). Maximum pressure policies in stochastic processing networks. *Operations Research*, *53*(2), 197–218.

[54] Dai, J. G., & Shi, P. (2019). Inpatient overflow: An approximate dynamic programming approach. *Manufacturing & Service Operations Management*, *21*(4), 894–911.

[55] Dall, T. M., Gallo, P. D., Chakrabarti, R., West, T., Semilla, A. P., & Storm, M. V. (2013). An aging population and growing disease burden will require a large and specialized health care workforce by 2025. *Health Affairs*, *32*(11), 2013–2020.

[56] Dattani, S., Rodés-Guirao, L., Ritchie, H., & Roser, M. (2023). *Mental Health*. Retrieved August 30, 2024, from https://ourworldindata.org/mental-health

[57] Dawande, M., Feng, Z., & Janakiraman, G. (2021). On the structure of bottlenecks in processes. *Management Science*, *67*(6), 3853–3870.

[58] de Boer, T. R. (2021). Simulation models to evaluate long term effects in the elderly care system in the Amsterdam region (MSc thesis). *Vrije Universiteit Amsterdam*.

[59] de Boer, T. R., Arntzen, R. J., Bekker, R., Buurman, B. M., Willems, H. C., & van der Mei, R. D. (2025). Process mining on national healthcare data for the discovery of patient journeys of older adults. *Journal of the American Medical Directors Association*, *26*(1), 105333.

[60] Decerle, J., Grunder, O., El Hassani, A. H., & Barakat, O. (2017). A general model for the home health care routing and scheduling problem with route balancing. *IFAC-PapersOnLine*, *50*(1), 14662–14667.

[61] Dickey, B., Cannon, N. L., McGuire, T. G., & Gudeman, J. E. (1986). The quarterway house: A two-year cost study of an experimental residential program. *Psychiatric Services*, *37*(11), 1136–1143.

[62] Dıaz, J. A., & Fernández, E. (2001). A Tabu search heuristic for the generalized assignment problem. *European Journal of Operational Research*, *132*(1), 22–38.

[63] European Commission. (2021). *Long-term care report: Trends, challenges and opportunities in an ageing society, volume II*.

[64] Evans, C. J., Potts, L., Dalrymple, U., Pring, A., Verne, J., Higginson, I. J., Gao, W., & SPACE. (2021). Characteristics and mortality rates among patients requiring intermediate care: A national cohort study using linked databases. *BMC Medicine*, *19*(48), 1–13.

[65] Fairweather, G. W., Sanders, D. H., Cressler, D. L., & Maynard, H. (2017). *Community life for the mentally ill: An alternative to institutional care*. Routledge.

[66] Farmer, E. M., Wagner, H. R., Burns, B. J., & Murray, M. (2016). Who goes where? Exploring factors related to placement among group homes. *Journal of Emotional and Behavioral Disorders*, *24*(1), 54–63.

[67] Gans, N., Koole, G., & Mandelbaum, A. (2003). Telephone call centers: Tutorial, review, and research prospects. *Manufacturing & Service Operations Management*, *5*(2), 79–141.

[68] Global News. (2021). *'Impossibly long': LTC waitlists, age discrimination among concerns in request for public inquiry.* Retrieved August 30, 2024, from https://%20globalnews.ca/news/7716972/waitlists-age-discrimination-concerns-public-in%20quiry-ltc/

[69] Goldstein, J. M., Dziobek, J. F., Clark, R., & Bassuk, E. L. (1990). Supportive housing for the chronically mentally ill matching clients with community environments. *The Journal of Nervous and Mental Disease*, *178*(7), 415–422.

[70] Government of South Africa, Department of Social Development. (2010). Audit of residential facilities.

[71] Green, L. (1980). A queueing system in which customers require a random number of servers. *Operations Research*, *28*(6), 1335–1346.

[72] Green, L. (1981). Comparing operating characteristics of queues in which customers require a random number of servers. *Management Science*, *27*(1), 65–74.

[73] Gualandi, R., Masella, C., Viglione, D., & Tartaglini, D. (2019). Exploring the hospital patient journey: What does the patient experience? *PloS One*, *14*(12), e0224899.

[74] Gurvich, I., & Whitt, W. (2009). Scheduling flexible servers with convex delay costs in many-server service systems. *Manufacturing & Service Operations Management*, *11*(2), 237–253.

[75] Gurvich, I., & Whitt, W. (2010). Service-level differentiation in many-server service systems via queue-ratio routing. *Operations Research*, *58*(2), 316–328.

[76] Gurvich, I., & Van Mieghem, J. A. (2015). Collaboration and multitasking in networks: Architectures, bottlenecks, and capacity. *Manufacturing & Service Operations Management*, *17*(1), 16–33.

[77] Gurvich, I., & Van Mieghem, J. A. (2018). Collaboration and multitasking in networks: Prioritization and achievable capacity. *Management Science*, *64*(5), 2390–2406.

[78] Harper, A., Pitt, M., De Prez, M., Dumlu, Z. Ö., Vasilakis, C., Forte, P., & Wood, R. (2021). A demand and capacity model for home-based intermediate care: Optimizing the 'step down' pathway. *2021 Winter Simulation Conference*, 1–12.

[79] Hasan, I., Bahalkeh, E., & Yih, Y. (2020). Evaluating intensive care unit admission and discharge policies using a discrete event simulation model. *Simulation*, *96*(6), 501–518.

[80] Heckman, G. A., Tannenbaum, C., Costa, A. P., Harkness, K., & McKelvie, R. S. (2014). The journey of the frail older adult with heart failure: Implications for management and health care systems. *Reviews in Clinical Gerontology*, *24*(4), 269–289.

[81] Heiner, K., Wallace, W. A., & Young, K. (1981). A resource allocation and evaluation model for providing services to the mentally retarded. *Management Science*, *27*(7), 769–784.

[82] Hinde, S., Bojke, L., Richardson, G., Birks, Y., Whittaker, W., Wilberforce, M., & Clegg, A. (2021). Delayed transfers of care for older people: A wider perspective. *Age and Ageing*, *50*(4), 1073–1076.

[83] Hirsch, J., Nicola, G., McGinty, G., Liu, R., Barr, R., Chittle, M., & Manchikanti, L. (2016). Icd-10: History and context. *American Journal of Neuroradiology*, *37*(4), 596–599.

[84] Holterman, S., Lahr, M., Wynia, K., Hettinga, M., & Buskens, E. (2020). Integrated care for older adults: A struggle for sustained implementation in Northern Netherlands. *International Journal of Integrated Care*, *20*(3).

[85] Holtman, G., Luijendijk, D., Havinga, H., & Zuidema, S. (2017). Kernmerken en beloop van patiënten in Eerstelijnsverblijf (in Dutch). *Tijdschrift voor Ouderengeneeskunde*.

[86] IkWoonLeefZorg. (2021). *Op de wachtlijst voor het verpleeghuis? Zo werkt het (in Dutch)*. Retrieved August 30, 2024, from https://www.ikwoonleefzorg.nl/hulp-en-zorg/op-de-wachtlijst-voor-het-verpleeghuis-zo-werkt-het

[87] Ilinca, S., & Calciolari, S. (2015). The patterns of health care utilization by elderly Europeans: Frailty and its implications for health systems. *Health Services Research*, *50*(1), 305–320.

[88] Inzitari, M., Gual, N., Roig, T., Colprim, D., Pérez-Bocanegra, C., San-José, A., & Jimenez, X. (2015). Geriatric screening tools to select older adults susceptible for direct transfer from the emergency department to subacute intermediate-care hospitalization. *Journal of the American Medical Directors Association*, *16*(10), 837–841.

[89] ISPH. (2021). *About the international society for priorities in health*. Retrieved April 15, 2021, from https://prioritiesinhealth.org/isph

[90] Jonckheere, M., & Borst, S. (2006). Stability of multi-class queueing systems with state-dependent service rates. *Valuetools'06*.

[91] Jordan, W. C., & Graves, S. C. (1995). Principles on the benefits of manufacturing process flexibility. *Management Science*, *41*(4), 577–594.

[92] Katsaliaki, K., Brailsford, S., Browning, D., & Knight, P. (2005). Mapping care pathways for the elderly. *Journal of Health Organization and Management*, *19*(1), 57–72.

[93] Kaye, H. S., Harrington, C., & LaPlante, M. P. (2010). Long-term care: Who gets it, who provides it, who pays, and how much? *Health Affairs*, *29*(1), 11–21.

[94] Kerimov, S., Ashlagi, I., & Gurvich, I. (2023). On the optimality of greedy policies in dynamic matching. *Operations Research*.

[95] Khatutsky, G., Ormond, C., Wiener, J. M., Greene, A. M., Johnson, R., Jessup, E. A. J., Vreeland, E., Sengupta, M., Caffrey, C., & Harris-Kojetin, L. D. (2016). Residential care communities and their residents in 2010: A national portrait. *DHHS publication*, (2016-1041).

[96] Kim, R. K., & Dymond, S. K. (2012). A national study of community living: Impact of type of residence and hours of in-home support. *Research and Practice for Persons with Severe Disabilities*, *37*(2), 116–129.

[97] Kim, R. K., & Dymond, S. K. (2020). What skills are critical for living in supported apartments and small group homes? *Journal of Developmental and Physical Disabilities*, *32*, 665–681.

[98] Knezevic, A. (2008). Overlapping confidence intervals and statistical significance. *StatNews: Cornell University Statistical Consulting Unit*, *73*(1).

[99] Koeleman, P. M., Bhulai, S., & van Meersbergen, M. (2012). Optimal patient and personnel scheduling policies for care-at-home service facilities. *European Journal of Operational Research*, *219*(3), 557–563.

[100] Koole, G. (1998). Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems*, *30*, 323–339.

[101] Koopmans, R. T., Lavrijsen, J. C., Hoek, J. F., Went, P. B., & Schols, J. M. (2010). Dutch elderly care physician: A new generation of nursing home physician specialists. *Journal of the American Geriatrics Society*, *58*(9), 1807–1809.

[102] Kydd, A. (2008). The patient experience of being a delayed discharge. *Journal of Nursing Management*, *16*, 121–126.

[103] Leff, H. S., Dada, M., & Graves, S. C. (1986). An LP planning model for a mental health community support system. *Management Science*, *32*(2), 139–155.

[104] Leff, J., & Szmidla, A. (2002). Evaluation of a special rehabilitation programme for patients who are difficult to place. *Social Psychiatry and Psychiatric Epidemiology*, *37*, 532–536.

[105] Legros, B., & Jouini, O. (2019). On the scheduling of operations in a chat contact center. *European Journal of Operational Research*, *274*(1), 303–316.

[106] Lim Fat, G. J., Gopaul, A., Pananos, A. D., & Taabazuing, M.-M. (2022). Healthcare-associated adverse events in alternate level of care patients awaiting long-term care in hospital. *Geriatrics*, *7*(4), 81.

[107] *LimeSurvey (in Dutch)*. (no date). Retrieved August 30, 2024, from https://www.limesurvey.org/nl

[108] Maarse, J. A. M., & Jeurissen, P. P. (2016). The policy and politics of the 2015 long-term care reform in the Netherlands. *Health Policy*, *120*(3), 241–245.

[109] Mandelbaum, A., & Stolyar, A. L. (2004). Scheduling flexible servers with convex delay costs: Heavy-traffic optimality of the generalized cμ-rule. *Operations Research*, *52*(6), 836–855.

[110] Mans, R. S., Schonenberg, M., Song, M., van der Aalst, W., & Bakker, P. J. (2009). Application of process mining in healthcare – a case study in a Dutch hospital. *Biomedical Engineering Systems and Technologies: International Joint Conference, BIOSTEC 2008*, 425–438.

[111] Marshall, A. H., McClean, S. I., & Millard, P. H. (2004). Addressing bed costs for the elderly: A new methodology for modelling patient outcomes and length of stay. *Health Care Management Science*, *7*, 27–33.

[112] Martin, G. P., Hewitt, G. J., Faulkner, T. A., & Parker, H. (2007). The organisation, form and function of intermediate care services and systems in England:

Results from a national survey. *Health & Social Care in the Community*, *15*(2), 146–154.

[113] Mary Martini, E., Garrett, N., Lindquist, T., & Isham, G. J. (2007). The boomers are coming: a total cost of care model of the impact of population aging on health care costs in the United States by major practice category. *Health Services Research*, *42*(1), 201–218.

[114] McCarthy, J., & Nelson, G. (2009). An evaluation of supportive housing: Qualitative and quantitative perspectives. *Canadian Journal of Community Mental Health*, *12*(1), 157–175.

[115] McCarthy, M. L., Zeger, S. L., Ding, R., Aronsky, D., Hoot, N. R., & Kelen, G. D. (2008). The challenge of predicting demand for emergency department services. *Academic Emergency Medicine*, *15*(4), 337–346.

[116] Meiland, F. J. M., Danse, J. A. C., Wendte, J. F., Gunning-Schepers, L. J., & Klazinga, N. S. (2002). Urgency coding as a dynamic tool in management of waiting lists for psychogeriatric nursing home care in the Netherlands. *Health Policy*, *60*, 171–184.

[117] Meiland, F. J. M., Danse, J. A. C., Wendte, J. F., Klazinga, N. S., & Gunning-Schepers, L. J. (2001). Caring for relatives with dementia - Caregiver experiences of relatives of patients on the waiting list for admission to a psychogeriatric nursing home in the Netherlands. *Scandinavian Journal of Public Health*, *29*, 113–121.

[118] Melis, R. J., Rikkert, M. G. O., Parker, S. G., & van Eijken, M. I. (2004). What is intermediate care?

[119] Mitzenmacher, M. (2001). The power of two choices in randomized load balancing. *IEEE Transactions on Parallel and Distributed Systems*, *12*(10), 1094–1104.

[120] Moynihan, R., Sanders, S., Michaleff, Z. A., Scott, A. M., Clark, J., To, E. J., Jones, M., Kitchener, E., Fox, M., Johansson, M., Lang, E., Duggan, A., Scott, I., & Albarqouni, L. (2021). Impact of COVID-19 pandemic on utilisation of healthcare services: a systematic review. *BMJ open*, *11*(3), e045343.

[121] Mukherjee, D., Borst, S. C., Van Leeuwaarden, J. S., & Whiting, P. A. (2018). Universality of power-of-d load balancing in many-server systems. *Stochastic Systems*, *8*(4), 265–292.

[122] Nancarrow, S., Moran, A., Freeman, J., Enderby, P., Dixon, S., et al. (2009). Looking inside the black box of community rehabilitation and intermediate care teams in the United Kingdom: An audit of service and staffing.

[123] Nederlandse Zorgautoriteit. (2022). Beleidsregel Eerstelijnsverblijf (in Dutch).

[124] Nederlandse Zorgautoriteit. (2024a). *DBC-zorgproducten (in Dutch)*. Retrieved August 30, 2024, from https://www.opendisdata.nl/msz/zorgproduct

[125] Nederlandse Zorgautoriteit. (2024b). *Kerncijfers wijkverpleging (in Dutch)*. Retrieved August 30, 2024, from https://www.nza.nl/zorgsectoren/wijkverpleging/kerncijfers-wijkverpleging

[126] Nelson, G., Hall, G. B., & Walsh-Bowers, R. (1997). A comparative evaluation of supportive apartments, group homes, and board-and-care homes for psychiatric consumer/survivors. *Journal of Community Psychology*, *25*(2), 167–188.

[127] Nino-Mora, J. (2009). Stochastic scheduling. In *Encyclopedia of optimization* (Pages 367–372). Springer.

[128] Nomura, K., Yamori, K., Takahashi, E., Miyoshi, T., & Tanaka, Y. (2001). Waiting time versus utility to download images. *4th Asia Pacific-Symposium on Information and Telecommunication Technologies*, 128–132.

[129] Noorain, S., Paola Scaparra, M., & Kotiadis, K. (2023). Mind the gap: a review of optimisation in mental healthcare service delivery. *Health Systems*, *12*(2), 133–166.

[130] Norwegian Institute of Public Health. (no date). *International language for drug utilization research: ATC/DDD*. Retrieved August 30, 2024, from https://atcddd.fhi.no/

[131] NOS. (no date). *Patiënten en zorgverleners dagen Staat en verzekeraars om lange wachtlijsten ggz (in Dutch)*. Retrieved August 30, 2024, from https://nos.nl/artikel/2522419-patienten-en-zorgverleners-dagen-staat-en-verzekeraars-om-lange-wachtlijsten-ggz

[132] NOS. (2020). *Kamer ontstemd: wachtlijsten verpleeghuizen flink langer dan verwacht (in Dutch)*. Retrieved August 30, 2024, from https://nos.nl/artikel/2318759-kamer-ontstemd-wachtlijsten-verpleeghuizen-flink-langer-dan-verwacht.html

[133] NOS. (2022). *Onderzoek: tekort aan zorgpersoneel op lange termijn alleen maar groter (in Dutch)*. Retrieved August 30, 2024, from https://nos.nl/artikel/2413851-onderzoek-tekort-aan-zorgpersoneel-op-lange-termijn-alleen-maar-groter

[134] OECD. (2019). *Health at a glance 2019*.

[135] Özkan, E. (2021). On the optimal control of parallel processing networks with resource collaboration and multitasking. *Stochastic Systems*, *11*(13), 284–303.

[136] Pagel, C., Richards, D. A., & Utley, M. (2012). A mathematical modelling approach for systems where the servers are almost always busy. *Computational and Mathematical Methods in Medicine*, *2012*(1), 290360.

[137] *Patiëntenfederatie Nederland (in Dutch)*. (no date). Retrieved August 30, 2024, from https://www.patientenfederatie.nl/

[138] Patrick, J. (2011). Access to long-term care: The true cause of hospital congestion? *Production and Operations Management*, *20*(3), 347–358.

[139] Pedlar, D., & Walker, J. (2004). The Overseas Service Veteran (OSV) at home pilot: How choice of care may impact use of nursing home beds and waiting lists. *Canadian Journal on Aging*, *23*(4), 367–369.

[140] Pellico-López, A., Fernández-Feito, A., Cantarero, D., Herrero-Montes, M., Cayón-de Las Cuevas, J., Parás-Bravo, P., & Paz-Zulueta, M. (2022). Cost of stay and characteristics of patients with stroke and delayed discharge for non-clinical reasons. *Scientific Reports*, *12*(1), 10854.

[141] Reichardt, L. A., Aarden, J. J., van Seben, R., van der Schaaf, M., Engelbert, R. H., Bosch, J. A., & Buurman, B. M. (2016). Unravelling the potential mechanisms behind hospitalization associated disability in older patients; the Hospital-Associated Disability and impact on daily Life (Hospital-ADL) cohort study protocol. *BMC Geriatrics*, *16*(1), 1–11.

[142] Reiman, M. I. (1991). A critically loaded multiclass Erlang loss system. *Queueing Systems*, *9*, 65–81.

[143] Ribbe, M. W. (1993). Care for the elderly: The role of the nursing home in the Dutch health care system. *International Psychogeriatrics*, *5*(2), 213–222.

[144] Rijksoverheid. (2022). Monitor ouderenzorg: Trendanalyses (deel 2) (in Dutch).

[145] Rijksoverheid. (2024a). *Wet langdurige zorg (in Dutch)*. Retrieved July 7, 2024, from https://www.rijksoverheid.nl/onderwerpen/langdurige-zorg

[146] Rijksoverheid. (2024b). *Wet maatschappelijke ondersteuning (in Dutch)*. Retrieved July 7, 2024, from https://www.rijksoverheid.nl/onderwerpen/maatschappelijke-ondersteuning

[147] Rijksoverheid. (2024c). *Zorgverzekeringswet (in Dutch)*. Retrieved July 7, 2024, from https://www.rijksoverheid.nl/onderwerpen/zorgverzekering

[148] RIVM. (2018). *Uitleg zorgzwaartepakket verpleging en verzorging (in Dutch)*. Retrieved August 30, 2024, from https://www.rivm.nl/sites/default/files/2018-11/ZZP-score.pdf

[149] RIVM. (2022). *Effecten van vergrijzing steeds beter zichtbaar in uitgaven gezondheidszorg (in Dutch)*. Retrieved August 30, 2024, from https://www.rivm.nl/nieuws/effecten-van-vergrijzing-steeds-beter-zichtbaar-in-uitgaven-gezondheidszorg

[150] Rockwood, K., & Mitnitski, A. (2011). Frailty defined by deficit accumulation and geriatric medicine defined by frailty. *Clinics in Geriatric Medicine*, *27*(1), 17–26.

[151] Saghafian, S., Kilinc, D., & Traub, S. (2022). *Dynamic assignment of patients to primary and secondary inpatient units: Is patience a virtue?*

[152] Saxena, S., Lora, A., Van Ommeren, M., Barrett, T., Morris, J., & Saraceno, B. (2007). WHO's assessment instrument for mental health systems: collecting essential information for policy and service delivery. *Psychiatric Services*, *58*(6), 816–821.

[153] Schimmelpfeng, K., Helber, S., & Kasper, S. (2012). Decision support for rehabilitation hospital scheduling. *OR Spectrum*, *34*(2), 461–489.

[154] Scott, J., Dawson, P., Heavey, E., De Brún, A., Buttery, A., Waring, J., & Flynn, D. (2021). Content analysis of patient safety incident reports for older adult patient transfers, handovers, and discharges: Do they serve organizations, staff, or patients? *Journal of Patient Safety*, *17*(8), e1744–e1758.

[155] Sezgin, D., O'Caoimh, R., O'Donovan, M. R., Salem, M. A., Kennelly, S., Samaniego, L. L., Carda, C. A., Rodriguez-Acuña, R., Inzitari, M., Hammar, T., et al. (2020). Defining the characteristics of intermediate care models including transitional care: An international Delphi study. *Aging Clinical and Experimental Research*, *32*, 2399–2410.

[156] Shapiro, E., Tate, R. B., & Tabisz, E. (1992). Waiting times for nursing-home placement: The impact of patients' choices. *Canadian Medical Association Journal*, *146*(8), 1343–1348.

[157] Sharma, A., Minh Duc, N. T., Luu Lam Thang, T., Nam, N. H., Ng, S. J., Abbas, K. S., Huy, N. T., Marušić, A., Paul, C. L., Kwok, J., Janette Karbwang, de Waure, C., Drummond, F. J., Kizawa, Y., Taal, E., Vermeulen, J., Lee,

G. H. M., Gyedu, A., Gia To, K., Verra, M. L., … Karamouzian, M. (2021). A Consensus-Based Checklist for Reporting of Survey Studies (CROSS). *Journal of General Internal Medicine*, *36*(10), 3179–3187.

[158] Shugarman, L. R., & Brown, J. A. (2006). Nursing home selection: How do consumers choose? Volume I: Findings from focus groups of consumers and information intermediaries.

[159] SIGRA. (2019). Krakende ketens in de zorg voor kwetsbare ouderen (in Dutch).

[160] SIGRA. (2024). *Samen voor betere zorg en welzijn (in Dutch)*. Retrieved August 30, 2024, from https://www.sigra.nl/

[161] Singh, D. A. (2014). *Effective management of long-term care facilities*. Jones & Bartlett Publishers, Sudbury, MA.

[162] Sinreich, D., & Marmor, Y. (2005). Emergency department operations: The basis for developing a simulation tool. *IIE transactions*, *37*(3), 233–245.

[163] Skipr. (2018). *Capaciteitsproblemen blijven kort herstel achtervolgen (in Dutch)*. Retrieved August 30, 2024, from https://www.skipr.nl/nieuws/capaciteitsproblemen-blijven-kort-herstel-achtervolgen/

[164] Spasova, S., Baeten, R., Vanhercke, B., et al. (2018). Challenges in long-term care in Europe. *Eurohealth*, *24*(4), 7–12.

[165] Staatscourant van het Koninkrijk der Nederlanden. (2016). Subsidieregeling Eerstelijnsverblijf (in Dutch).

[166] Taherdoost, H. (2017). Determining sample size; how to calculate survey sample size. *International Journal of Economics and Management Systems*, *2*.

[167] TMI. (2024). *Specialist ouderengeneeskunde (in Dutch)*. Retrieved August 30, 2024, from https://www.tmi.nl/artsen/specialist-ouderengeneeskunde/#:~:text=Een%20specialist%20ouderengeneeskunde%20houdt%20zich,alleen%20een%20arts%20voor%20ouderen

[168] Trebble, T. M., Hansi, N., Hydes, T., Smith, M. A., & Baker, M. (2010). Process mapping the patient journey: An introduction. *BMJ*, *341*, c4078–c4078.

[169] United Nations Department of Economic and Social Affairs. (2022). World population prospects 2022: Summary of results.

[170] Valtorta, N. K., Moore, D. C., Barron, L., Stow, D., & Hanratty, B. (2018). Older adults' social relationships and health care utilization: A systematic review. *American Journal of Public Health*, *108*(4), e1–e10.

[171] van den Besselaar, J. H., Arntzen, R. J., MacNeil Vroomen, J. L., Hertogh, C. M. P. M., & Buurman, B. M. (2023). Short term residential care in the Netherlands: Patient and facility characteristics from a national database and survey. *Submitted*.

[172] van den Besselaar, J. H., Hartel, L., Wammes, J. D., MacNeil-Vroomen, J. L., & Buurman, B. M. (2021). 'Patients come with two garbage bags full of problems and we have to sort them.' A qualitative study of the experiences of healthcare professionals on patients admitted to short-term residential care in the Netherlands. *Age and Ageing*, *50*(4), 1361–1370.

[173] van der Aalst, W. (2011). Process mining: discovering and improving spaghetti and lasagna processes. *2011 IEEE symposium on computational intelligence and data mining*, 1–7.

[174] van der Aalst, W. (2012). Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems*, *3*(2), 1–17.

[175] van der Aalst, W. (2016). *Process mining*. Springer.

[176] van der Perk, K. (2023). Capacity allocation for the reduction of bottlenecks in the Amsterdam elderly care system (MSc thesis. *Vrije Universiteit Amsterdam*.

[177] Vandenbrouckel, J. P., von Elm, E., Altman, D. G., Gotzsche, P. C., Mulrow, C. D., Pocock, S. J., Poole, C., Schlesselman, J. J., & Egger, M. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Medicine*, *4*(10), 1628–1655.

[178] Vektis. (2018a). *1 op de 11 65-plussers heeft een ongeplande ziekenhuisopname (in Dutch)*. Retrieved August 30, 2024, from https://www.vektis.nl/intelligence/publicaties/1-op-de-11-65-plussers-heeft-een-ongeplande-ziekenhuisopname

[179] Vektis. (2018b). *Vooral vrouwen van 80+ maken gebruik van Eerstelijnsverblijf (in Dutch)*. Retrieved August 30, 2024, from https://www.vektis.nl/actueel/vooral-vrouwen-van-80-maken-gebruiken-van-eerstelijns-verblijf#:~:text=Sinds%201%20januari%202017%20wordt,er%20gebruikmaken%20van%20het%20ELV.

[180] Vektis. (2018c). Zorgthermometer: Inzicht in de ouderenzorg (in Dutch).

[181] Vektis. (2020). Ligduur Eerstelijnsverblijf loopt terug, uitstroom naar huis neemt toe (in Dutch).

[182] Vektis. (2021). Factsheet Geriatrische Revalidatiezorg (in Dutch).

[183] Velasquez, J. S., & McCubbin, H. (1980). Towards establishing the effectiveness of community-based residential treatment. *Journal of Social Service Research*, *3*(4), 337–359.

[184] *Verenso: vereniging van specialisten ouderengeneeskunde (in Dutch)*. (no date). Retrieved August 30, 2024, from https://www.verenso.nl/

[185] Verhaegh, K. J., MacNeil-Vroomen, J. L., Eslami, S., Geerlings, S. E., de Rooij, S. E., & Buurman, B. M. (2014). Transitional care interventions prevent hospital readmissions for adults with chronic illnesses. *Health Affairs*, *33*(9), 1531–1539.

[186] Vrije Universiteit. (no date). *Lang wachten op geestelijke gezondheidszorg verkleint de kans op een baan (in Dutch)*. Retrieved August 30, 2024, from https://vu.nl/nl/nieuws/2023/lang-wachten-op-geestelijke-gezondheidszorg-verkleint-de-kans-op-een-baan

[187] Wallace, R. B., & Whitt, W. (2005). A staffing algorithm for call centers with skill-based routing. *Manufacturing & Service Operations Management*, *7*(4), 276–294.

[188] Walsh, B., Fogg, C., Harris, S., Roderick, P., de Lusignan, S., England, T., Clegg, A., Brailsford, S., & Fraser, S. D. (2023). Frailty transitions and prevalence in an ageing population: longitudinal analysis of primary care data from an open cohort of adults aged 50 and over in England, 2006–2017. *Age and Ageing*, *52*(5), afad058.

[189] Wang, J., Mao, Y., McGarry, B., & Temkin-Greener, H. (2022). Post-acute care transitions and outcomes among medicare beneficiaries in assisted living communities. *Journal of the American Geriatrics Society*, *70*(5), 1429–1441.

[190] *Wijziging Wet langdurige zorg, Wet maatschappelijke ondersteuning 2015, Jeugdwet en Zorgverzekeringswet (34.279)*. (no date). Retrieved March 13, 2024, from https://www.eerstekamer.nl/wetsvoorstel/34279_wijziging_wet_langdurige

[191] Williams, E., Gartner, D., & Harper, P. (2021). A survey of OR/MS models on care planning for frail and elderly patients. *Operations Research for Health Care*, *31*, 100325.

[192] Williams, M. E., Latta, J., & Conversano, P. (2008). Eliminating the wait for mental health services. *The Journal of Behavioral Health Services & Research*, *35*, 107–114.

[193] Wolpert, J., & Wolpert, E. R. (1976). The relocation of released mental hospital patients into residential communities. *Policy Sciences*, *7*(1), 31–51.

[194] World Health Organization. (2001). Mental health in Europe: Country reports from the WHO European network on mental health.

[195] Yoo, S., Cho, M., Kim, E., Kim, S., Sim, Y., Yoo, D., Hwang, H., & Song, M. (2016). Assessment of hospital processes using a process mining technique: Outpatient process analysis at a tertiary hospital. *International Journal of Medical Informatics*, *88*, 34–43.

[196] Young, J., Gladman, J. R., Forsyth, D. R., & Holditch, C. (2015). The second national audit of intermediate care. *Age and Ageing*, *44*(2), 182–184.

[197] Young, J., Hulme, C., Smith, A., Buckell, J., Godfrey, M., Holditch, C., Grantham, J., Tucker, H., Enderby, P., Gladman, J., et al. (2020). Measuring and optimising the efficiency of community hospital inpatient care for older people: The MoCHA mixed-methods study. *Health Services and Delivery Research*, *8*(1).

[198] Zenios, S. A., Chertow, G. M., & Wein, L. M. (2000). Dynamic allocation of kidneys to candidates on the transplant waiting list. *Operations Research*, *48*(4), 549–569.

[199] Zorginstituut Nederland. (2024). *Kosten van de kortdurende zorg stijgen harder dan aantal gebruikers (in Dutch)*. Retrieved August 30, 2024, from https://www.zorgcijfersdatabank.nl/nieuws/factsheet-kortdurende-zorg

[200] *Zorginstituut Nederland (in Dutch)*. (no date). Retrieved August 30, 2024, from https://www.zorginstituutnederland.nl/

[201] Zorgverzekeraars Nederland. (2020). Landelijke rapportage wachtlijsten Verpleging en Verzorging (in Dutch).

[202] Zychlinski, N., Chan, C. W., & Dong, J. (2023). Managing queues with different resource requirements. *Operations Research*, *71*(4), 1387–1413.

Summary

This thesis is about the healthcare system for older adults from a mathematical and analytical perspective. Below we give a brief summary of each of the chapters.

In Chapter 2 we examine healthcare utilization patterns by Process Mining among 3,177,203 older adults in the Netherlands from 2017 to 2019, aiming to uncover common patient journeys. We found that 44% of these individuals experienced one or more patient journeys, resulting in a total of 2,469,663 journeys. Most journeys were simple and short, with 95% involving four or fewer care transitions. For the 85+ population, 90% of journeys also had four or fewer transitions. Long-term care forms, such as home care, primary care, and long-term care, dominated the time spent within the healthcare system. The study reveals that while the majority of older adults have straightforward healthcare needs, a smaller subset requires more complex and prolonged care, particularly among those aged 85 and older. Reducing the number of transitions in this group, though affecting fewer individuals, could have a significant impact on the overall healthcare system.

Chapter 3 examines the characteristics of older adults admitted to Short-Term Residential Care (STRC) in the Netherlands and compared organizational differences between STRC facilities. Patient profiles were analyzed using a national STRC database from 2018 to 2019, revealing that 68,682 older adults were admitted, with a significant proportion being female, living alone, and prescribed multiple medications. Additionally, an email survey of 176 STRC facilities provided insights into organizational variations, such as ward locations, bed numbers, and care delivery practices. Results showed that 30.1% of facilities offered care in independent wards, while others were integrated within geriatric or long-term care wards. The study found that most facilities admitted patients during evenings, nights, or weekends and employed registered nurses and paramedics. The high mortality rate observed within 24 months of STRC admission underscores the importance of developing care pathways that address reablement, diagnostics, and palliative care for older adults.

Chapter 4 investigates the prolonged waiting times for Intermediate Care (IC) in the Netherlands, which often lead to unnecessary hospital admissions. Using data from older adults who received IC in Amsterdam in 2019, we analyzed patient inflows, outflows, and characteristics to understand the system's bottlenecks. A process map of the main pathways into and out of IC was created, followed by the development of a Discrete Event Simulation (DES) model. Our analysis revealed that waiting times are primarily due to inefficient triage and application processes, rather than a lack of bed capacity. The median waiting time of 1.8 days often leads to hospitalization. Through sensitivity analysis, we found that streamlining the application process and allowing evening and weekend admissions could significantly reduce unwanted hospitalizations. This study highlights the need for data-driven approaches to optimize IC logistics, providing a foundation for effective policy decisions aimed at improving patient care.

Chapter 5 addresses the inefficiencies in Long-Term Care (LTC) systems that lead to excessive waiting times and patient abandonments, particularly in the context of an aging population. These delays are not due to a lack of total capacity, but result from systematic inefficiencies in matching patients to geographically dispersed care centers. To solve this, we propose a novel allocation method that optimally balances waiting times with individual patient preferences and flexibility. Using an MDP model, we developed an optimal placement policy for patients-in-need. Our results show that for small instances, the mean optimality gap of the allocation model is just 1.3%, indicating a highly efficient placement process. Moreover, our allocation model better serves individual preferences, offering a promising approach to creating patient-centered and sustainable LTC solutions that can effectively meet the increasing demand for care in aging populations.

Chapter 6 addresses the challenge of long waiting times for nursing homes by implementing a preference-based allocation model for older adults. The model's performance was evaluated through simulations across three case studies in the Netherlands, covering urban areas like Amsterdam and Rotterdam, as well as the rural region of Twente. Data on nursing homes and capacities were collected, and preference profiles were developed to capture waiting time preferences and flexibility. The simulation compared the current allocation practices with the proposed model, focusing on outcomes like waiting times and alignment with preferences. Results showed that the allocation model significantly reduced waiting times — by at least a factor of two in Rotterdam and Twente — and more older adults were placed in their preferred nursing homes. The study concludes that the model outperforms existing waiting-line policies, is easy to implement, and has the potential to be extended to other care settings, demonstrating the value of mathematical models in optimizing care for older adults.

Chapter 7 tackles the issue of matching persons with severe mental illness to appropriate residential facilities, where current placement policies often disadvantage harder-to-place clients by giving priority to easier placements. To address this imbalance, we propose the so-called Knapsack-Based Routing (KBR) model, which aims to efficiently allocate clients to server pools while considering both resource constraints and individual preferences. Our approach utilizes a split-horizon method that incorporates integer quadratic programming to account for both immediate and long-term scheduling impacts. Through extensive simulation experiments, the split-horizon method demonstrated superior performance over

benchmark instances in specific cases of KBR, including scenarios involving overflow and strategic idleness. The method achieved a mean optimality gap of 2.3% in the developed benchmark instances. When applied to the allocation process for mental health facilities in Amsterdam, the method significantly improved waiting times, creating a more equitable distribution for both easy-to-place and hard-to-place clients.

Portfolio

Conference presentations

Conference	Location	Date
INFORMS Annual Meeting	Anaheim (virtual)	September 2021
Beta Conference	Soesterberg	November 2021
LNMB	Lunteren (virtual)	January 2022
Geriatrics Days	Breda (virtual)	February 2022
INFORMS Annual Meeting	Indianapolis	October 2022
Science Days on Geriatrics*	Utrecht	November 2022
Workshop on Queueing Theory	Obergurgl	December 2022
International Conference on Integrated Care	Antwerp	May 2023
INFORMS Healthcare	Toronto	July 2023
Workshop on Queueing Theory	Obergurgl	December 2023
StochMod	Milan	June 2024

* Won the prize for best scientific presentation.

Teaching and supervision

Supervision of master thesis Computational Science of Herman Aadamsoo (2022)
Supervision of master thesis Business Analytics of Tim de Boer (2022)
Supervision of master thesis Business Analytics of Jimmy Gijssel (2022)
Teaching Project OBP for the master Business Analytics (2022)
Supervision of master thesis Business Analytics of Katja van der Perk (2023)

About the Author

Rebekka was born on April 10, 1994, at the Amsterdam UMC location AMC and was raised in Amsterdam Bijlmer. After finishing high school in Amsterdam Zuid, she obtained a bachelor's degree in Econometrics and Operations Research from Rijksuniversiteit Groningen. In addition to her studies, she worked as a student assistant and exam trainer, which made her fall in love with teaching.

Motivated by this newfound interest, Rebekka decided to pursue a master's degree in Science Education & Communication from Utrecht University, which included a first-grade mathematics teaching certificate. She completed her master's thesis internship at Cito, in the department that develops high school mathematics exams.

After obtaining her master's degree, Rebekka wanted to learn more about Operations Research, so she pursued a second master's degree in Econometrics and Management Science from Erasmus University Rotterdam. She wrote her master's thesis at PostNL under the supervision of Thije van Barneveld. Thije is a former PhD student of Rob van der Mei, which led to the next step in her career: obtaining a PhD degree at Centrum Wiskunde & Informatica. For one year during her PhD, she worked part-time as a mathematics teacher at Christelijk Gymnasium Utrecht.

Rebekka has an older brother, Siebe, with a severe intellectual disability and autism. Rebekka therefore understands the importance of good healthcare, the consequences of cutbacks, and the impact of excessive waiting times. She regards the fact that she was offered a PhD position that combined her interests in stochastic models and the healthcare system as a 'rare event', which she accepted gratefully.