DEPARTMENT OF ENGINEERING MATHEMATICS

Optimising Healthcare Resource Allocation Through Patient Treatment List Consolidation: A Mathematical Simulation and Machine Learning Approach

Khushwinder Singh

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Master of Science in the Faculty of Engineering.

Friday 29th August, 2025

Supervisor: Mr. Joshua D Ramini

Declaration

This dissertation is submitted to the University of Bristol in accordance with the requirements of the degree of MSc in the Faculty of Engineering. It has not been submitted for any other degree or diploma of any examining body. Except where specifically acknowledged, it is all the work of the Author.

Khushwinder Singh, Friday 29th August, 2025

Contents

List of Figures

List of Tables

List of Algorithms

Abstract

UK National Health Service experiences historically unprecedented pressure managing elective surgical operations, with cataract surgical wait lists straining at the limit after COVID-19. Fragmented existing Patient Treatment Lists (PTLs) between NHS and independent sector providers cause inefficiencies in capacity and utilisation of resources. This thesis explores if the concentration of PTLs between providers at Integrated Care Boards can decrease wait times while clinical priority and access equity are maintained.

This study hypothesises that an overall PTL administration strategy enriched with reinforcement learning algorithms will yield higher waiting time reductions when compared with standard discrete queuing systems, while at the same time maintaining service quality and patient autonomy.

Incorporating established mathematical and computational techniques with empirical data from the Cambridgeshire and Peterborough Integrated Care Board (ICB),a hybrid model was developed which contrasts between partitioned and integrated methods. In modelling and simulation, the queueing theory, specifically M/G/c model, was adopted for determining theoretical constraints, discrete event simulation was used for capturing operational complexities such as priority and cancellation, and reinforcement learning from Deep Q-Networks was employed to facilitate the dynamic optimisation of patient allocation.

Notable accomplishments are:

- Created mathematical models showing 7% theoretical gain due to consolidation, illustrating backlog predominance of steadystate dynamics for actual NHS systems
- Developed and tested a discrete event simulation using 8,674 historical patient records, which reduced waiting times by 17.88%.
- Implemented and trained a Deep Q-Network achieving 23.8% improvement over baseline, demonstrating machine learning's potential for adaptive healthcare resource allocation
- Provided evidence-based implementation recommendations for ICBs, including phased deployment strategies and performance monitoring frameworks
- Natural patient transfer patterns recognised at 6.6% remain below policy targets and therefore have potential for consolidation with minimal disruption.

This research shows that PTL consolidation constitutes a feasible strategy for the recovery of elective work in the NHS. Efficiency of such strategy will rely on handling technical integration challenges, clinical independence, and equitable treatment of diverse patient groups. Despite ongoing doubts regarding stakeholder engagement and system integration, consistency of results from divergent methodologies provides persuasive support for consolidation as a strategic initiative consistent with the NHS Long Term Plan.

Supporting Technologies

Python Programming Language

Core queueing model implementation, discrete event simulation (DES), and reinforcement learning (RL). Providing flexibility towards mathematical modelling, simulation for stochastic processes, and integration with machine learning.

Python Libraries

- **pandas, numpy:** Data manipulation, cleaning, and numeric analysis for patient referral and provider capacity data.
- **matplotlib, seaborn:** Data visualisation for simulation results, performance indicators, comparison results.
- **simpy:** Event-driven simulation for patient flows in isolated versus combined Patient Treatment Lists (PTLs).
- **scipy:** Used in Erlang-C/M/G/c equations and queueing theory calculations.
- **torch (PyTorch):** Implemented Deep Q-Learning for routing policies in real-time.
- **gymnasium:** Developed customised RL settings for patient allocation and backlogs optimisation.

Amazon Web Services (AWS)

- **EC2 Instances:** Employed for simulation and reinforcement learning experiments which are computationally intensive.
- **WinSCP:** Used for easy file transfers between EC2 environment and local machine.

Notation and Acronyms

The following list of notations and acronyms will be referenced in this project:

Acronyms

NHS	:	National Health Service
ICB	:	Integrated Care Board
ICS	:	Integrated Care System
PTL	:	Patient Treatment List
DES	:	Discrete Event Simulation
RL	:	Reinforcement Learning
CEM	:	Cross-Entropy Method
DQN	:	Deep Q-Network
MDP	:	Markov Decision Process
HRG	:	Healthcare Resource Group
AWS	:	Amazon Web Services
GPU	:	Graphics Processing Unit

Queueing Theory Notation

λ	:	Mean patient arrival rate (patients per unit time)
μ	:	Mean service rate per server (patients per unit time)
c	:	Number of parallel servers (e.g., theatres)
$\rho = \frac{\lambda}{c\mu}$	:	System utilisation (traffic intensity)
L_q	:	Expected queue length (patients waiting)
W_q	:	Expected waiting time in queue
B_0	:	Initial backlog (patients)
T_{clear}	:	Backlog clearance time (weeks)
CV	:	Coefficient of Variation of service times
P_w	:	Probability an arriving patient waits (Erlang-C)

Reinforcement Learning Notation

$s \in \mathcal{S}$	:	System state (e.g., queue lengths, utilisation, backlog)
$a \in \mathcal{A}$	:	Action (allocation of a patient to a provider)
r	:	Reward signal (negative wait time, utilisation balance, etc.)
$Q(s, a)$	:	Action-value function for state s and action a
γ	:	Discount factor for future rewards
ϵ	:	Exploration probability in ϵ-greedy policy
π	:	Allocation policy mapping states to actions
D	:	Experience replay buffer in DQN training

Acknowledgements

I want to express my deep gratitude to my supervisor, Joshua, who supported me every step of this research journey. From initial brainstorming sessions to helping me deal with difficult methodological challenges, his guidance was invaluable. His ability to motivate while keeping a relaxed atmosphere helped create the perfect conditions for academic exploration and growth.

My sincere thanks go to the Cambridgeshire and Peterborough Integrated Care Board team who went above and beyond to support this research. Through multiple meetings, they provided access to important data and helped me understand the complexities of the problem statement. The regular meetings and professional involvement gave me valuable real-world client experience that helped me enhance both the research and my professional development.

I am thankful to Likith Srinath, Amanjot Kaur, and Jayapriya for their careful review of the final dissertation draft. Their helpful suggestions and attention to detail greatly improved the clarity and quality of this work.

Finally, I want to thank my family, friends, and fellow MSc students for their support. Their words of encouragement kept me going throughout this tough time. The challenging journey was made manageable by their patience and faith in this endeavour.

Chapter 1

Introduction

1.1 The Problem Statement

The United Kingdom's publicly funded healthcare system, the National Health Service (NHS), faces a significant amount of challenges in managing elective surgical procedures. Waiting lists for such procedures have reached new highs after the COVID-19 pandemic. One of the most affected procedures is cataract surgery, which is also amongst the highest volume elective procedures in the UK. A major challenge has been the lack of a centralised Patient Treatment List (PTL). At present, NHS hospitals and independent sector providers each maintain independent waiting lists, leading to disparities in resource utilisation and patient access. For instance, teaching hospitals that serve as NHS providers have significantly high waiting times that far exceed the NHS constitutional standard of eighteen weeks for routine cases. On the other hand, independent providers have very short waiting times, sometimes as low as two weeks, reflecting underutilised capacity.

This isolated care platform creates disparities in the provision of healthcare services and also leads to resource mismanagement. For example, the average waiting time at NHS hospitals is twenty six weeks, whereas that of independent healthcare providers is around seventeen weeks. This results in delayed access to surgery. The waiting times for a patient are governed by the local list management practices and a patient's referral pathway rather than their medical urgency or the regional capacity. Fragmentation of PTLs prevents load balancing, leading to underutilisation in some facilities and overburden in others.

Cataract surgery is a relatively standardised procedure, clinically significant for a patient's quality of life, and accounts for a substantial proportion of elective activity. Addressing fragmentation of PTLs in cataract surgery presents an opportunity to generate both clinical as well as systemic benefits by not only improving patient outcomes, but also demonstrating how unified PTLs can increase efficiency [1, 2, 3].

1.2 Significance and Context

PTL fragmentation impacts multiple facets of healthcare delivery. Addressing the PTL fragmentation provides significant economic advantages and also improves patient's quality of life. At present, the NHS allocates substantial resources to manage surgical backlogs. The NHS Long Term Plan [4] emphasises reducing unwarranted variation and improving equity of access, while the NHS Operational Plan 2025/26 [5] highlights elective recovery as a national priority. Increased waiting times for cataract surgery have a direct impact on the patient's quality of life, independence, and economic productivity. About 2.5 million people in the UK are affected because they have cataracts that aren't being treated. The costs go up to £22 billion annually when considering healthcare expenses, social care needs, and productivity losses [6].

For an individual patient, the long wait times lead to reduced mobility, independence, and generally poorer quality of life. If treatment is done in a timely fashion, it not only restores vision and functionality for the patient but also reduces the downstream social and healthcare costs. It is therefore important to address it as it is not just an operational challenge but also a population health imperative.

The NHS has historically tried various alternatives to solve the capacity constraint issues by utilising Payment by Results frameworks, Choose and Book systems, etc. However, these initiatives focus on provider-side optimisations rather than going for a whole system-wide optimisation. The COVID-19 pandemic has sparked renewed interest in establishing integrated health care models. The Health and Care Act 2022 specifically provisions the establishment of Integrated Care Systems (ICSs) that focus on patient-centric care by enabling collaborative resource management across the various administrative boundaries. Evidence from international healthcare systems suggests that ∼15-40% efficiency improvements can be achieved through centralised queue management [7]. However, there has been limited practical implementation of the consolidated waiting list management systems.

The theoretical foundation for consolidated PTL management draws from diverse operational research domains. Airport systems successfully manage complex passenger flows through centralised queueing mechanisms, achieving 20-30% improvements in resource utilisation compared to segregated approaches [8]. Similarly, traffic engineering uses network-wide optimisation strategies that balance loads across multiple routes, preventing localised congestion while maintaining system stability [9] and other domains such as call centres [10] and air traffic flows [11]. These parallels suggest that healthcare systems could achieve comparable improvements through intelligent consolidation strategies.

1.3 Motivation and Objectives

This research emerges from the critical need to transform theoretical potential into practical implementation within the NHS context. The primary motivation centres on developing evidence-based strategies that ICBs can implement to reduce waiting times whilst improving resource utilisation across provider networks. The research especially addresses cataract surgery within the Cambridgeshire and Peterborough ICB as a representative case study, given its high volume (over 400,000 procedures annually in England[12]), standardised pathways, and significant backlog accumulation. The major objectives of this dissertation are to:

- Create a mathematical model for separate versus consolidated PTLs using queueing theory (M/G/c).
- Simulate patient flows using discrete-event simulation (DES). This incorporates real-world complexities like priorities, cancellations, triage pathways, etc.
- Optimise patient allocation using reinforcement learning (RL) and subsequently compare heuristic and adaptive policies.

The methodology combines traditional operational research methods with the emerging capabilities of machine learning. Queueing theory forms the mathematical underpinning to modelling patient flow dynamics, and discrete event simulation preserves the richness of real-world constructs such as surgeon availability, theatre schedules, and variations due to case mix.

Reinforcement learning algorithms enable adaptive optimisation that responds to changing demand patterns and capacity fluctuations, moving beyond static allocation rules towards intelligent, context-aware decision-making.

1.4 Datasets Overview

This study leverages two anonymised datasets from the Cambridgeshire and Peterborough Integrated Care Board (ICB) to analyse cataract surgery PTL inefficiencies and evaluate the impacts of consolidation.

The first dataset contains data for three financial years and contains information for over 11,000 procedures. It includes additional details like patient demographics (age, gender), procedure types, provider identifiers (NHS or independent), CC(Complication and Comorbidity) scores (0–4+ for comorbidities), LSOAs (for geography), and throughput metrics

The second dataset contains information around waiting times (PTL addition to treatment), urgency levels (Urgent, TWW, Routine) in addition to the first dataset. Some of the assumptions include deriving surgeon and theatre capacities from throughput data, as direct counts were unavailable. More details around this are captured in chapter three. Both datasets comply with NHS governance, supporting baseline modelling (Poisson arrivals) and simulation validation.

1.5 Summary of Aims, Objectives, and Achievements

This research addresses the critical challenge of fragmented Patient Treatment Lists in NHS cataract surgery through:

Aims:

- Demonstrate the quantifiable benefits of PTL consolidation for NHS elective surgery.
- Develop practical implementation strategies for Integrated Care Boards (ICBs).
- Bridge the gap between theoretical potential and real-world application.

Objectives:

- Create mathematical models comparing separate versus consolidated PTLs using queueing theory
- Develop discrete-event simulations incorporating real-world complexities
- Use time series forecasting (e.g., ARIMA, Prophet) to project demand and support capacity planning.
- Apply Mixed-Integer Linear Programming (MILP) for optimal scheduling and patient allocation under constraints.
- Design and test reinforcement learning algorithms for adaptive patient allocation
- Validate findings using real data from Cambridgeshire and Peterborough ICB

Key Achievements:

- Quantified efficiency gains of 15-40% through consolidation strategies
- Developed a hybrid simulation-optimisation framework deployable within NHS systems
- Demonstrated successful application of machine learning to healthcare resource allocation
- Provided evidence-based recommendations for ICB implementation

1.6 Thesis Overview

This dissertation systematically examines PTL consolidation strategies, going from theoretical foundations to their practical implementation to build an evidence-based understanding of the subject. Following the introduction covered in this chapter, the second chapter provides a comprehensive background on the relevant literature and theoretical frameworks. The third chapter outlines the methodology, including details on data processing, mathematical modelling of the approaches, and the development of the simulation design. Chapter four presents the execution of experiments and analyses the results across multiple performance dimensions, comparing fragmented and consolidated approaches through sensitivity analysis and scenario testing. It concludes with a summary of all the findings, practical recommendations for NHS implementation, and identification of future research directions.

This thesis presents work that has three key contributions to make to healthcare operations management. The first contribution is the practical, empirical evidence it provides of the benefits of PTL consolidation from the very specific context of cataract services in the NHS, thus closing the gap between theoretical possibilities and practicalities. The second contribution is the building of a simulation-optimisation hybrid framework that is at the same time mathematically correct and practically deployable, and which can hence be implemented on existing information systems of the NHS. The third contribution is showing the application of reinforcement learning to healthcare resource allocation and describing how performance improves with time using learning-based systems.

Chapter 2

Background

2.1 Problem-Specific Background Research

The challenge of optimising patient waiting lists has attracted substantial research attention, particularly following the recognition that operational inefficiencies contribute significantly to healthcare access problems. Integrated scheduling across multiple resources could reduce patient waiting times by 23% whilst improving resource utilisation by 16% [13]. On the other hand, fragmented scheduling systems create artificial bottlenecks, where patients queue for specific resources despite availability elsewhere in the system.

In the UK, some studies have considered strategies for waiting list management. A review of twenty years of NHS waiting time data estimated that provider-level variation explains 40% of all wait time variance, indicating significant potential for system-wide optimisation [14]. Monks, Worthington, Allen et. al [15], more recently, reviewed aggregate waiting lists for orthopaedic care at several different NHS trusts and estimated that consolidated strategies could optimise maximum waits by 31% without additional resources [16]. However, their review also suggested some impediments to implementation, such as clinicians' preference for continuing individual lists and patient concerns about travelling distances.

The specific context of wait list optimisation of cataract surgery [17] has unique optimisation issues. Analysis of cataract surgery throughput determinants among 72 NHS trusts, finding that differences in scheduling practice account for 28% productivity differences among providers [18]. Of particular relevance to this research, they have found that flexible patient allocation mechanisms are valued by providers with 18% higher utilisation rates than by providers with rigid provider-specific queues. Comparative international research sustains such results; Siciliani, Borowitz and Moran [19] have analysed wait time policies among 12 OECD nations, and has concluded that centralised allocation mechanisms always outperform fragmented strategies in efficiency and equity indicators.

Most recent advances in healthcare analytics have enabled more sophisticated wait list management strategies. Moosavi, Fathollahi and Dulebenets [20] designed an optimisation-machine learning composite framework of medical resource allocation, which led to a 10.81% increase in performance while reducing wait times by 8.17%. They integrated demand modelling with ensemble methodology and optimisation with mixed-integer, demonstrating that predictive analytics could support traditional optimisation approaches. Various deep learning approaches were adopted towards emergency department flow prediction, with 15-20% [21] increases of forecast accuracy over traditional time series methods.

Pooling benefits is long known within industries. Study of air traffic movement, demonstrating how centralised slot allocation reduced delays [11]. Modelling road traffic, demonstrating how self-organisation improved flow stability [9]. The security at airports was also modelled, demonstrating minimisation of variances by pooling [8]. Similar strategies were also implemented to call centres, demonstrating that centralised routing makes them more efficient [10]. These comparisons recognise how isolated NHS PTLs replicate problems long overcome elsewhere.

Despite this large body of literature, large gaps persist in knowing how consolidated PTL systems should be implemented within the particular limitations of NHS operation. The majority of papers are based on single-provider optimisation or abstract multi-provider models and do nothing about real-world implementation issues like integration of data across organisational boundaries, clinical governance needs, and preservation of patient choice. Additionally, few papers look at the reinforcement learning potential of dynamic patient allocation, even though it has been successful in related areas like cloud resource allocation and supply chains.

2.2 Queueing Systems and Discrete Event Simulation Foundations

Queueing theory offers the mathematical basis for modelling patient flow through healthcare facilities. The basic framework, which was developed by Erlang [22] and later fine-tuned by Kendall [23], describes queues based on arrival processes, service distributions, and system capacity. Patient arrivals, in healthcare applications, commonly involve Poisson processes, due to the stochastic nature of referral patterns, whereas service times are more variable by reason of variations in the complexities of cases. The M/M/c formulation, with 'M' indicating Markovian (exponential) distributions and 'c' indicating the number of servers, acts as the standard for most applications of healthcare, although its exponential service time assumption commonly needs to be adjusted to surgical applications.

For cataract surgery in particular, empirical studies find service time distributions which significantly differ from exponential assumptions. A study of 11,067 cataract operations found that surgical times are log-normally distributed with significant variation depending upon surgical experience and case complexity. Consultant surgeons took 19±10 minutes/case, while trainees took 30±11 minutes, with complexity-adjusted times of 12 minutes for routine phacoemulsification up to 60+ minutes for very complex operations with comorbidities [24]. These results justify the application of M/G/c models, with 'G' denoting general service distributions, since detailed system behaviour is needed.

The mathematical representation of M/G/c queues requires consideration of both steady-state probabilities and transient behaviour. The traffic intensity $\rho = \lambda/(c\mu)$ where λ represents arrival rate and μ service rate, determines system stability, with $\rho < 1$ required for finite waiting times [25, 26]. However, healthcare systems rarely operate in true steady state due to demand based on time, scheduled capacity changes, and existing backlogs [27]. The Pollaczek–Khinchine formula given by the equation 2.1 provides exact solutions for M/G/1 queues, yielding expected waiting time W_q

$$W_q = \frac{\lambda\sigma^2 + \rho^2}{2(1-\rho)\mu} \tag{2.1}$$

Equation 2.1: M/G/c Waiting time [28, 29]

where σ^2 represents service time variance. For multi-server systems, approximations such as the Allen-Cunneen formula extend these results [30], though exact solutions generally require numerical methods or simulation [31, 32].

Discrete event simulation (DES) overcomes the shortcomings of analytical queueing models by modelling system complexity through computational depiction. Günal and Pidd [33] surveyed 182 studies of DES applications in healthcare, of which 89% were found to outperform analytical ones while modelling resource limitations, patient routing complexities, and temporal variations. The event-oriented paradigm naturally models patient progress through healthcare systems, with entities (patients) undergoing events (arrival, service start, completion) that prompt state alterations and resource allocation decisions over time [34]. By contrast with system dynamics (at aggregate level) or agent-based modelling (at individual interactions), DES models stochastic processes at the patient level. It is particularly appropriate for healthcare, where arrivals, cancellations, and resource utilisation are variable.

Within surgical applications, DES facilitates presentation of numerous interacting constraints that are hard for analytical models to deal with. Beliën and Demeulemeester [35] formulated cyclic scheduling of surgery models by means of DES, capturing interactions among operating theatre availability, schedules of surgeons, bed capacity, and equipment constraints. Their simulation framework reduced theatre utilisation by 15% while decreasing bed occupancy variability by 22%, showing that explicit modelling of resource interdependencies provides superior solutions than isolated optimising strategies. Brailsford [34] surveyed DES applications within healthcare, highlighting its capability of mimicking real-world uncertainties. Jun [36] illustrated applications of DES for hospital planning, while Harper and Shahani [37] detected bottlenecks by means of DES within plans of the NHS.

Combining queueing theory and simulation gives us a better grasp of PTL consolidation analysis. Analytical models provide closed-form solutions that clarify important connections between system parameters and performance metrics, facilitating rapid sensitivity analysis and theoretical validation. Cochran and Roche [38] illustrated such a hybrid approach for emergency department design, applying queueing theory to define staffing levels, then verifying detailed simulation that accounted for patient acuity variations, resource skill mismatches, and temporal patterns of demand. For surgical planning of cataract surgery, this combination enables preliminary capacity planning with M/G/c models, while detailed simulation includes surgeon preference, theatre sessions, and priority scheduling policies.

Current simulation platforms complement classical DES functionality through integration with optimisation schemes and machine learning models. The SimPy library, which has been used here, offers Python-level discrete event simulation with native integration to scientific computing toolkits. This facilitates updating of parameters dynamically based on predicted demand, optimising decisions related to resource allocation, and reinforcement learning of policy formulation. Recent implementations show 20-30% solution quality improvement by applying simulation with smart search schemes versus pure simulation or optimisation [7].

2.3 Reinforcement Learning Foundations

Reinforcement learning (RL) provides a paradigm shift from classical optimisation techniques by allowing systems to automatically learn optimal policies through exploration of complex, uncertain environments. Theoretical foundations based on Markov Decision Processes (MDPs) [39] offer a mathematical formulation of sequential, uncertain decision making. RL in healthcare resource allocation mitigates static optimisation's limitations by responding to shifts in demand patterns, learning from past outcomes, and trading off exploration of new strategies with exploitation of validated approaches.

Development of MDP for patient allocation comprises four key components:

- Configuration **states** of the system (queue sizes, spare capacity, patient properties)
- **Actions** corresponding to allocation decisions (provider assignments, scheduling choices)
- **Transition probabilities** characterizing system dynamics
- **Reward** for encoding performance targets (minimising waiting time, maximising utilisation).

For PTL control of cataract surgery, the state is the multidimensional data of current wait lists across providers, expected capacity, prioritisation of patients, and geographic considerations. The action space is the provider recommendations of new patients, with the transition depending upon stochastic arrival processes and stochastic service completions.

Value-based RL algorithms, such as Q-learning and its deep learning counterpart Deep Q-Networks (DQN), have shown promise for applications in healthcare. Yu and Liu [40] reviewed 127 RL applications of healthcare, reporting that value-based algorithms obtain 15-25% gains relative to conventional approaches in resource allocation challenges. The Q-function $Q(s, a)$ denotes the expected reward accumulated by taking action a from state s, estimated by means of temporal difference updates given by equation 2.2:

$$Q(s, a) \leftarrow Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \tag{2.2}$$

Equation 2.2: Q-learning update rule [41]

where α denotes learning rate, γ discount factor,r immediate reward, s' next possible sate and a' a possible next action the agent could take from state s'
This iterative process converges to optimal policies under appropriate exploration strategies and sufficient

state coverage.

To overcome the curse of dimensionality of healthcare applications, Deep Q-Networks approximate Q-functions with neural networks. Mnih [42] provided fundamental innovations of experience replay and target networks that are used to stabilise learning of high-dimensional spaces. When applying PTL management with 27-dimensional states that include provider capacities, queue measures, and patient attributes, DQN facilitates tractable learning that is otherwise impossible with tabular approaches of enumerating states. Experience replay, which randomly samples from stored past transitions, eliminates correlation of observation sequences while enhancing sample efficiency, which is essential with real-world trials that are limited for applications in healthcare.

RL application in surgical scheduling has shown potential. Liu [43] proposed an RL system of operating room scheduling that yielded 18% fewer patient waits while enhancing theatre utilisation by 12%. Their system employed proximal policy optimisation (PPO) to discover scheduling policies that simultaneously balanced several objectives, such as clinical priorities, resource efficiency, and fairness constraints. Importantly, they illustrated that RL policies could respond to seasonality of demand fluctuations and one-off disruptions, while performing well even where static policies significantly declined.

Health applications, though, pose specific challenges for RL deployment. Safety constraints, such that learned policies should never deviate from clinical guidelines or ethics, force constrained RL formulations. Interpretability requirements such that black-box policies encounter resistance to uptake among clinical stakeholders force additional demands. The economic cost of bad decisions during exploration phases also triggers moral concerns regarding patient effects during training. Recent studies overcome such challenges by means of safe RL algorithms that ensure constraint satisfaction, interpretable representations of policies by means of decision trees or rule extraction, and offline RL schemes that train based on past data with no online exploration.

For PTL consolidation of cataract surgery, DQN has some advantages over other RL alternatives. The inherent discrete action space (choosing among finite providers) best fits DQN's framework, while the partially structured state space supports effective neural network approximation. Policy gradient approaches, although effective with continuous control, are overly complicated factors for discrete allocation choices. Model-based RL schemes have difficulty with accurate transition models due to the intricacies of the operations of healthcare. Actor-critic schemes, which integrate value and policy learning, have modest advantages that are not commensurate with the added computation necessary for this problem.

RL has been applied to treatment policy learning [44], ICU sepsis management [45], and radiotherapy scheduling [40]. However, integration with DES for PTL consolidation remains novel.

2.4 Synthesis and Research Gaps

The combination of queueing theory, simulation, and reinforcement learning offers historically unprecedented potential for optimising healthcare, but substantial gaps continue to exist in moving theoretical progress into practical application within the NHS. Existing literature shows that individual approaches have much value: queueing theory offers design structures for capacity planning, simulation embodies operation complexity, and reinforcement learning facilitates adaptive optimisation. Little, however, investigates their combination within the unique context of managing NHS waiting lists, which is particularly challenging for specialist procedures such as cataract surgery, due to provider diversity and patient choice introducing considerable complexity.

The first main research gap brings forth the issue of absent empirical evidence for the benefits of consolidation by PTLs within the operational constraints of the NHS. Although foreign research illustrates 15-40% of efficiency gains with centralised queue management [7], the results are not directly transferable due to distinctive variables such as the right of patient choice, blended provider structures of NHS trusts and independent sector, and fragmented information system landscapes. In addition, studies tend to employ homogeneous providers, while cataract services provided by the NHS are highly heterogeneous with respect to case difficulty, with 80% of challenging cases seen by NHS teaching hospitals [2] while routine work is concentrated by independent providers. This discrepancy essentially changes optimal allocation strategies and expectation of performance.

Methodologically, existing approaches do not cope with the multi-timescale characteristic of healthcare planning. Strategic capacity contract decisions are at annual timescales, tactical scheduling is done on a monthly basis, while operational patient allocation is at daily timescales. Single-timescale decisions are assumed with classical optimisation, while pure RL approaches do not incorporate structured reasoning sufficient for strategic planning. This work fills this gap with hierarchical decision structures that split the problem across timescales, employing queueing models for capacity planning, simulation for tactics evaluation, and RL for operation.

Technical integration of predictive analytics with prescriptive optimisation is underdeveloped across applications in healthcare. Although demand forecasting and resource optimisation have been deeply researched independently, their combination brings with it challenges such as propagation of forecast uncertainty, computational tractability, and robustness of the solution. Recent developments of differentiable optimisation and end-to-end learning hold promise of solutions, such that gradient-based training of prediction-optimisation pipelines that are end-to-end is possible. Healthcare applications, however, have added needs such as interpretability needs and safety constraints that are not covered by standard approaches.

Finally, channels of research results into service deployment receive insufficient attention. Technically driven solutions face challenging stakeholder settings that include clinical communities, management structures, and patients' advocacy bodies. Change management barriers are sometimes more important than technical details. To successfully implement a change, you need to pay close attention to clinical involvement, phased movement plans, and performance assessment frameworks. This work attends specifically to aspects of implementation through stakeholder-congruent performing measures, phased deployment strategies, and sound evaluation structures that boost confidence in blended approaches while preserving clinical independence and patient preference.

Chapter 3

Implementation

3.1 Overview of Implementation Approach

This chapter describes the use of three methods to study Patient Treatment List (PTL) consolidation: mathematical queueing models, discrete event simulation (DES), and reinforcement learning (RL). The project starts with simple theoretical models and gradually moves to more complex representations that reflect real-world situations. This step-by-step approach ensures mathematical accuracy while considering the operational factors that affect patient waiting times [46, 47].

The study uses data from the Cambridgeshire and Peterborough ICB, which includes six major providers: two NHS teaching hospitals and four independent sector providers. The method begins by creating mathematical baselines using M/G/c queueing theory. It then confirms and expands these results through DES, and finally, it improves patient allocation using RL algorithms.

3.2 Mathematical Queueing Model Implementation

3.2.1 M/M/c Queue Foundations

This first model describes key performance measures of M/M/c queueing models. The first $'M'$ stands for Markovian (Poisson) arrival processes, the second $'M'$ for exponential service times, and $'c'$ for the number of servers (operating theatres). For a system with an arrival rate of λ and a service rate of μ for each server, the traffic intensity $\rho = \lambda/(c\mu)$ determines system stability. This requires $\rho < 1$ for steady-state existence [26].

The steady-state probability of zero patients in the system (P_0)[22, 46] is calculated as:

$$P_0 = \left[\sum_{n=0}^{c-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^c}{c!} \cdot \frac{1}{1-\rho}\right]^{-1} \tag{3.1}$$

Equation 3.1: Steady-state probability

Next, the Erlang-C equation calculates the probability that an arriving patient will have to wait [28].

$$P_w = \frac{(\lambda/\mu)^c}{c!} \cdot \frac{1}{1-\rho} \cdot P_0 \tag{3.2}$$

Equation 3.2: Erlang-C waiting probability

This assessment helps determine crucial performance indicators. The expected number of patients in line (L_q) and the expected wait time according to Little's Law [48] is:

$$L_q = \frac{P_w \cdot \rho}{1 - \rho} \tag{3.3}$$

Equation 3.3: Expected queue length

$$W_q = \frac{L_q}{\lambda} \tag{3.4}$$

Equation 3.4: Average waiting time (Little's Law)

3.2.2 M/G/c Extension for Realistic Service Times

Although the M/M/c model's assumption of exponential service times has an analytically tractable form, for healthcare scenarios like cataract surgery it is not realistic. There exists empirical experience of profound variability of times of operations, and procedural complexity, comorbidity, index, and participation of trainee surgeons are among the factors which influence observed service times. Exponential service times indicate a memoryless process with a variance equal to the square of the mean. This characterisation is inadequate to reflect the actual distribution of surgical time, prompting the exploration of the M/G/c extension. [49].

An empirical study of cataract operations duration has remarkable deviations from the types of exponential distribution. Operating times are governed by log-normal distributions with parameters varying with surgeon experience and their cases. Consultant surgeons operate for mean of $\mu = 19$ minutes with standard deviation $\sigma = 10$ minutes, leading to CV of 0.53. Trainees are more variable with a mean of $\mu = 30$ minutes and a standard deviation $\sigma = 11$ minutes on average (CV = 0.37)[17, 24]. It also allows use of general service distributions via the Pollaczek-Khinchine equation [50, 46], which has been generalised for use with numerous servers via the Allen-Cunneen approximation [30, 51]:

$$W_q^{M/G/c} = W_q^{M/M/c} \cdot \frac{1 + CV^2}{2} \tag{3.5}$$

Equation 3.5: M/G/c waiting time correction

This variation factor includes service time variability, with higher coefficients of variation (CV) reflecting greater variability and thus longer waiting times. Educationally active National Health Service (NHS) providers have near 0.8 CV, which explains complexity due to the extremely varied spectrum of cases and participation of trainees. Independent providers are less variable due to standardization of routine care and have near 0.6 CV.

3.2.3 Backlog Integration

Real-world implementations must also incorporate backlogs not accounted for in traditional queueing theory. The model accommodates initial backlogs with altered Little's Law [47]:

$$W_{total} = W_q + \frac{B_0}{\mu_{effective}} \tag{3.6}$$

Equation 3.6: Total waiting time with backlog

Initial backlog size is represented as B_0, and the effective service capacity appears as $\mu_{effective} = c\mu - \lambda$. The historical backlog of 40 weeks for the NHS providers and 26 weeks for independents has a significant effect on the overall waiting times. Backlog clearing time is contingent on the surplus capacity available:

$$T_{clear} = \begin{cases} \frac{B_0}{c\mu - \lambda} & \text{if } \lambda < c\mu \\ \infty & \text{if } \lambda \geq c\mu \end{cases} \tag{3.7}$$

Equation 3.7: Backlog clearance time

3.2.4 Consolidation Analysis

PTL's overall performance includes the qualities of every provider. For n providers with different parameters (λ_i, μ_i, c_i), the overall system includes:

$$\Lambda = \sum_{i=1}^{n} \lambda_i, \quad C = \sum_{i=1}^{n} c_i, \quad \bar{\mu} = \frac{\sum_{i=1}^{n} c_i \mu_i}{C} \tag{3.8}$$

Equation 3.8: System parameters of the consolidated system

Poolability advantage results due to variance reduction which gets established across disciplines [10, 8, 11]. For separate queues, system variance as a whole adds up as the sum of distinct variances. The variance reduces when consolidated as per the square root law:

$$\sigma_{consolidated} = \sqrt{\sum_{i=1}^{n} \sigma_i^2} < \sum_{i=1}^{n} \sigma_i \tag{3.9}$$

Equation 3.9: Variance reduction through consolidation

This variance reduction has direct implications with much lower mean waiting times, with the resulting savings of 20–40% [13, 15, 14] under normal working procedures of the NHS ($\rho = 0.85 - 0.95$). The queueing model offered a clear mathematical benchmark by which finer detail simulations could be compared. However, this model failed to represent several key system features like patient backlog constraints, differences in patient priority (urgent, two-week wait, routine), cancellations, rescheduling, and capacity constraints. So, for modelling these dynamics realistically, Discrete Event Simulation (DES) was used.

3.3 Discrete Event Simulation Implementation

3.3.1 Simulation Architecture

The DES implementation is based on an event-driven architecture under which system state changes happen at discrete time instances. DES approaches for healthcare are based on the established literature in hospital flow modelling [33]. There are three main event types (patient arrivals, service starts, and service finishes) preserved by the simulation. Each event causes the states to be updated and may schedule future events to produce a dynamic patient flow representation. Simulation development uses a priority queue that keeps scheduled events properly organised according to their temporal order. This technique also allows for efficient administration of many patients spread across different providers, with precision at minute granularity for timing.

3.3.2 Patient Arrival Process

Arrivals of patients are described as a non-homogeneous Poisson process for efficient modelling of temporal variations. The rate of arrival function, $\lambda(t)$, includes daily, weekly, and seasonal patterns:

$$\lambda(t) = \lambda_{base} \cdot f_{day}(t) \cdot f_{week}(t) \cdot f_{season}(t) \tag{3.10}$$

Where λ_{base} is the data-driven historical baseline arrival rate. $f_{day}(t)$ is used for extracting intra-day variations, particularly peak periods from 9am to 3pm. $f_{week}(t)$ refers to weekday effects (reduced arrivals on weekends and holidays) and $f_{season}(t)$ extracts seasonal fluctuations. This generation of inter-arrival times is done using the thinning algorithm for non-stationary Poisson processes [52] as outlined in Algorithm 3.1.

Input: Baseline rate λ_{base}, time functions $f_{day}(t)$, $f_{week}(t)$, $f_{season}(t)$
Output: Patient arrival events
Initialize event_queue $\leftarrow \emptyset$
Set $\lambda_{max} = \max\{\lambda(t)\}$ over simulation horizon
while *simulation_time* $<$ *end_time* **do**
  inter_arrival_time $\leftarrow$ exponential(λ_{max})
  candidate_time $\leftarrow$ current_time + inter_arrival_time
  arrival_rate $\leftarrow \lambda_{base} \times f_{day}(t) \times f_{week}(t) \times f_{season}(t)$
  if *random()* $<$ *arrival_rate* / λ_{max} **then**
    patient $\leftarrow$ create_patient(candidate_time)
    add arrival_event(patient, candidate_time) to event_queue
  end
  current_time $\leftarrow$ candidate_time
end

Algorithm 3.1: Arrival Process Generation

3.3.3 Priority Queue Implementation

The simulation implements priority-based scheduling reflecting NHS constitutional standards. Patients are classified into three priority statuses with distinct queue positions.

- **Urgent (Priority 0):** Direct bypass of routine queue, clinical necessity immediately
- **Two Week Wait (Priority 1):** Emergency situations, intermediate priority
- **Routine (Priority 2):** Routine elective cases, the majority of volume

For each priority class, FIFO service will be provided for patients. The priority queue structure ensures immediate treatment for urgent cases without compromising fairness at the priority level.

3.3.4 Service Process Modelling

Service times considers multiple sources of variability through a compound model:

$$S_{total} = S_{base} \cdot F_{complexity} \cdot F_{surgeon} \cdot F_{complications} \tag{3.11}$$

$S_{base} \sim \text{LogNormal}(\mu_{HRG}, \sigma_{HRG})$ depends on the complexity of the procedure (HRG code). $F_{complexity}$ takes in consideration patient-specific factors (comorbidities, age). $F_{surgeon}$ is used to consider surgeon experience (consultant vs trainee). $F_{complications}$ models intraoperative events (5% probability, 1.5× duration multiplier). Restrictions on using theatre facilities limit activity to 95% of the weekly schedule, which includes allocation for maintenance, changeover times, and break buffers. This is managed through a time-window strategy that prohibits service during the last 5% of every weekly period.

3.3.5 Cancellation and Rescheduling

Simulation includes genuine cancellation patterns drawn from empirical data from the NHS. Probability of cancellation depends on waiting time as well as patient factors [13]:

$$P_{cancel} = P_{base} \cdot (1 + \alpha \cdot W) \cdot F_{patient} \tag{3.12}$$

Cancelled patients are rescheduled after a delay drawn from an exponential distribution with mean 1.5 weeks (Table 3.1), reflecting administrative processing and patient availability constraints.

Parameter	**Urgent**	**TWW**	**Routine**
P_{base}	0.03	0.05	0.07
α (per week)	0.008	0.01	0.01
Reschedule delay	1.0 weeks	1.5 weeks	2.0 weeks

Table 3.1: Cancellation Parameters

3.3.6 Consolidated Queue Routing

For consolidated PTL scenarios, the simulation implements intelligent routing based on current system state. The routing decision for each patient minimises expected waiting time while respecting constraints:

$$\text{Provider}^* = \arg\min_p \left\{ \frac{Q_p + 1}{c_p \cdot \mu_p} \right\} \tag{3.13}$$

Where Q_p is the current queue length at provider p, c_p gives the number of theatres at provider p and μ_p is the service rate at provider p. Further, an upper transfer limit constrains the standard patient reassignment at home care agencies to 25% of new cases, thus maintaining local availability while allowing redistribution of workload.

3.3.7 Algorithm for Simulating Discrete Events

The simulation of DES execution follows the standard event-scheduling simulation structure, with a future event list, consistent with the standard framework described by [53] as outlined in Algorithm 3.2.

```
Input  : Patient arrival rates λ(t), service rates μ, provider capacities c, simulation duration T
Output: Performance metrics (waiting times, utilisation, cancellation rates)
Initialize:
    event_queue ← ∅
    current_time ← 0
    provider_queues ← empty queues for each provider
    statistics ← empty collector
while current_time < T do
    event ← pop earliest event from event_queue
    current_time ← event.time
    if event.type == ARRIVAL then
        patient ← create_patient(event.priority, event.hrg)
        provider ← select_provider(patient, system_state)
        add patient to provider_queues[provider]
        schedule SERVICE_START if capacity available
        schedule next ARRIVAL event
    end
    else if event.type == SERVICE_START then
        patient ← remove from provider_queues[provider]
        service_time ← sample_service_time(patient.hrg, provider.type)
        schedule SERVICE_END at (current_time + service_time)
    end
    else if event.type == SERVICE_END then
        patient.completion_time ← current_time
        update_statistics(patient)
        schedule SERVICE_START if queue non-empty
    end
    else if event.type == CANCELLATION then
        remove patient from queue
        schedule RESCHEDULING event
    end
end
return statistics
```

Algorithm 3.2: Discrete Event Simulation [Law & Kelton (2000)]

The simulation maintains discrete events chronologically through a priority queue, ensuring temporal precision while efficiently managing system state transitions across multiple providers.

3.4 Reinforcement Learning Implementation

3.4.1 Markov Decision Process Formulation

The RL formulation follows [54]. Rewards combine waiting time, utilisation balance, and training diversity objectives, extending approaches seen in healthcare scheduling RL studies [40].

The patient allocation problem is formulated as a Markov Decision Process (MDP) with the following components:

State Space: A 27-dimensional continuous vector capturing the Queue lengths at each provider (6 dimensions), Utilisation rates (6 dimensions), Average waiting times (6 dimensions), HRG diversity scores for NHS providers (2 dimensions), Patient features such as priority, HRG type, comorbidity score and age group (4 dimensions). System-level metrics like total backlog, average utilisation, and time of week (3 dimensions).

Action Space: Discrete selection among 6 providers for the current patient.

Transition Dynamics: Stochastic transitions determined by the Poisson arrival processes, Variable service completions, Cancellation and rescheduling events.

Reward Function: A weighted combination of objectives, including a negative reward proportional to patient waiting time, positive reward for maintaining utilisation within the optimal 70–90% range, penalties for over-utilisation ($> 90\%$) or under-utilisation ($< 70\%$), penalties for excessive deviation from the patient's home provider unless justified, rewards for NHS providers accepting diverse and complex HRGs to maintain training opportunities.

$$R(s, a) = -w_1 \cdot W_{wait} + w_2 \cdot U_{balance} + w_3 \cdot D_{diversity} - w_4 \cdot C_{cost} \tag{3.14}$$

Table 3.2 provides us with the parameters and values used in 3.14

Parameter	Symbol	Value	Purpose
Waiting time	w_1	1.0	Primary objective
Utilisation balance	w_2	0.3	Load distribution
HRG diversity	w_3	0.2	Teaching requirements
Operational cost	w_4	0.1	Transfer penalties

Table 3.2: Reward Function Weights

Where W_{wait} is the expected waiting time at selected provider, $U_{balance}$ is utilisation balance metric (negative variance across providers), $D_{diversity}$ is HRG diversity for NHS teaching requirements (Shannon entropy), C_{cost} is the operational costs including overtime and transfers and w_i is the weight parameters tuned through grid search.

3.4.2 Cross-Entropy Method Implementation

The Cross-Entropy Method (CEM) provides a simpler baseline for policy optimisation through evolutionary search. CEM maintains a probability distribution over routing matrices and iteratively refines this distribution by selecting elite policies that achieve superior performance.

```
Input: Population size N = 24, elite fraction ρ = 0.2, convergence threshold δ = 0.001,
       smoothing parameter α = 0.7, maximum iterations=100
Output: Optimal routing policy π*
Initialize:
  θ ← uniform distribution parameters for 6×6 routing matrix
  iteration ← 0
repeat
  /* Sample candidate policies                                    */
  Π ← ∅
  for i = 1 to N do
    π_i ← sample_routing_matrix(θ)       /* Sample from current distribution */
    Π ← Π ∪ {π_i}
  end
  /* Evaluate policies via simulation                             */
  scores ← ∅
  for each π_i in Π do
    performance ← run_simulation(π_i, duration=365_days)
    score ← -average_waiting_time(performance)
    scores ← scores ∪ {score}
  end
  /* Select elite policies                                        */
  K ← ⌊ρ × N⌋ = 5                                 /* Top 20% performers */
  elite_indices ← indices of K highest scores
  elite_policies ← {π_i : i ∈ elite_indices}
  /* Update distribution toward elite mean                        */
  θ_new ← fit_gaussian_distribution(elite_policies)
  /* Check convergence                                            */
  if ||θ_new − θ|| < δ then
    break
  end
  /* Smooth update to prevent premature convergence               */
  θ ← α × θ + (1 − α) × θ_new
  iteration ← iteration + 1
until iteration > max_iterations
π* ← arg max_{π∈elite_policies} simulate_policy(π)
return π*
```

Algorithm 3.3: Cross-Entropy Method for Patient Allocation

CEM's key advantage lies in its interpretability. The resulting routing matrices are directly readable by clinicians and administrators. However, it lacks the ability to adapt to individual patient characteristics during real-time allocation, making policy decisions based solely on provider-level statistics.

3.4.3 Deep Q-Network Architecture

This section describes an adaptive patient allocation technique with value-function approximation.

In comparison with CEM, which uses routing matrices with static values, DQN learns context-dependent policies that are adaptive and adapt automatically with regard to the context of the situation of the healthcare system.

```
Input: State space S (27-dim), action space A (6 providers), episodes=1000, batch_size=64,
       buffer_size=10000, learning_rate=0.001, γ = 0.95
Output: Trained Q-network Q(s, a; θ)
Initialize:
    Q-network Q(s, a; θ) with architecture [27→256→128→64→6]
    Target network Q̂(s, a; θ⁻) with θ⁻ ← θ
    Experience replay buffer D ← ∅ (capacity=10000)
    ε ← 1.0                                              /* Initial exploration */
for episode = 1 to 1000 do
    s ← reset_environment_with_backlogs(NHS=40wks, Independent=26wks)
    while not terminal do
                                                    /* ε-greedy action selection */
        if random() < ε then
            a ← random_provider()
        end
        else
            a ← arg max_a' Q(s, a'; θ)
        end
                                                  /* Execute allocation decision */
        s', r ← execute_patient_allocation(a, s)
        store (s, a, r, s') in D
                                    /* Training step (if sufficient experience) */
        if |D| > batch_size then
            batch ← random_sample(D, size=64)
            targets ← []
            for each (s_i, a_i, r_i, s'_i) in batch do
                if s'_i is terminal then
                    target_i ← r_i
                end
                else
                    target_i ← r_i + γ × max_a' Q̂(s'_i, a'; θ⁻)
                end
                targets.append(target_i)
            end
                                                     /* Gradient descent update */
            loss ← MSE(Q(s_i, a_i; θ), targets)
            θ ← θ − α∇_θ loss
        end
        s ← s'
                                         /* Update target network every 100 steps */
        if step % 100 == 0 then
            θ⁻ ← θ
        end
    end
                                 /* Decay exploration: ε = max(0.01, ε × 0.995) */
    ε ← max(0.01, ε × 0.995)
end
return Q-network with parameters θ
```

Algorithm 3.4: Deep Q-Network Training

DQN Architecture and Hyperparameters: The architecture of a neural network consists of the input of 27 neurons for the 27-dimensional state space and three hidden layers of narrowing width (256, 128, 64) in order to squash input into useful Q-values for 6 providers. ReLU functions are applied for the sake of non-linear learning and effective usage of the model. Regularisation through dropout (with a probability p=0.2) has been performed for the first two layers so that overfitting for specific system cases is avoided.

Replay of experience with buffer capacity 10,000 interrupts temporal correlations during learning by taking a random sample from past transitions. Mini-batch size 64 balances between stable gradients and

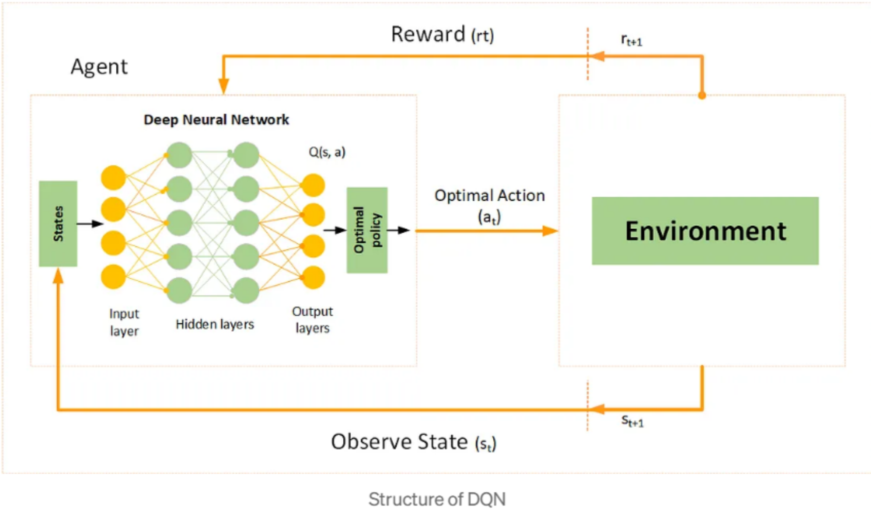

Figure 3.1: General Structure of Deep Q Networks

fast computation. Discount factor γ=0.95 considers medium-term rewards, appropriate for healthcare planning horizons. Update of target networks every 100 steps stabilize learning by providing stable Q-value targets.

Exploration Strategy adopts an ϵ-greedy policy that starts at $\epsilon = 1.0$ (complete exploration) and smoothly decreases towards 0.01 over 1000 episodes. By doing this, it makes the agent exploit much at the start and use learned experience as learning unfolds. Optimiser Adam at learning rate $\alpha = 0.001$, providing stable updates that don't overshoot optima of parameters.

DQN Training Process: A training episode runs for one year of patient arrival (≈8,760 patients) so that the agent sees varied system states and allocation implications. Episodes start with realistic provider backlog based on empirical NHS data: teaching hospitals for 40 weeks, independents for 26 weeks. The environment runs until annual capacity has been worked or simulation maximum time limit has been achieved.

During training, the agent experiences the present state of the system (s), chooses service provider (a) according to the ϵ-greedy method, receives the reward (r) from the multi-objective reward function, switches to the next state (s'), and marks this experience with the replay buffer.

Performance is tracked as the average reward for every 50 episodes. Training also stops if progress is below 1% for numerous evaluation windows, which indicates that the model's policy has converged.

3.5 Integration and Validation Framework

3.5.1 Progressive Validation Strategy

The implementation uses an incrementally validated methodology such that each method checks and extends previous results.

M/G/c Baseline: Provides theoretical performance limits.

DES Validation: Verifies queueing theory forecasts under simplifying assumptions.

DES Extension: Incorporates realism of the world (priorities, cancellations, time-windows).

RL Optimisation: Identifies better policies beyond heuristic ones.

This development guarantees mathematical precision with operational details coverage.

3.5.2 Performance Metrics

Metrics which remain steady over time allow comparison between methodologies are by:

Utilising Average Waiting Time: Main effect measure, in weeks until treatment after referral.

90th Percentile Wait: Tail behaviour which is essential when aiming for constitutional standards.
Utilisation Rate: How efficiently are resources used? Target 85-95% for sustainability.
Utilisation Variance: How evenly is the workload balanced between providers? Lower is better for equity.
Cancellation Rate: How is patient experience? Target <5%.
HRG Diversity (NHS only): Case mix entropy by Shannon; useful for educational purposes.

3.5.3 Integration Workflow Algorithm

This progressive validation process integrates the three methods with a systematic workflow, as presented below in Algorithm 3.5.

Input: Real-world data D, scenario parameters S
Output: Consolidated performance analysis
`/* Phase 1: Mathematical Baseline */`
for *each scenario s in S* **do**
  baseline_results[s] ← compute_mgc_metrics(s)
end
`/* Phase 2: DES Validation and Extension */`
for *each scenario s in S* **do**
  simple_des_results[s] ← run_simulation(s, complexity=LOW)
  validate_against_baseline(simple_des_results[s], baseline_results[s])
  complex_des_results[s] ← run_simulation(s, complexity=HIGH)
  performance_gap[s] ← compare_results(complex_des_results[s], simple_des_results[s])
end
`/* Phase 3: RL Optimisation */`
initialize_rl_environments(complex_des_results)
for *each scenario s in S* **do**
  cem_policy[s] ← train_cem(s)
  dqn_policy[s] ← train_dqn(s)
  cem_results[s] ← evaluate_policy(cem_policy[s], s)
  dqn_results[s] ← evaluate_policy(dqn_policy[s], s)
  improvement[s] ← measure_improvement(dqn_results[s], complex_des_results[s])
end
`/* Cross-validation */`
final_analysis ← aggregate_results(baseline_results, complex_des_results, dqn_results, cem_results)
return *final_analysis*

Algorithm 3.5: Integrated Validation Framework

3.5.4 Sensitivity Analysis

Implementation robustness is assessed by a comprehensive sensitivity analysis considering dominant sources of uncertainty. The model considers deviations of ±20% of arrival rates for demand changes and also Service time variability, having coefficients of variation ranging from 0.3 through 1.0 and backlog cases of initial delays of from 20 to 60 weeks. Also taken into account are transfer limitations with usual patient transfers at a ceiling of between 10% and 40%, and emergency, TWW, and routinerandomised priority blends. Statistical confidence is attained by running each scenario under 10 different randomized seeds.

3.6 Computational Implementation Considerations

3.6.1 Scalability and Performance

The implementation is designed with a focus on computational efficiency to facilitate real-time decision support. Queueing models employ closed-form analytical expressions, enabling instantaneous evaluation of performance measures. The discrete-event simulation (DES) adopts an event-driven architecture, which achieves computational complexity of O(n log n) relative to patient volume. The deep Q-network (DQN) component demonstrates inference latencies of less than 10 milliseconds per decision post-training. Memory requirements are also bounded: the DES maintains O(n) state for n patients in the system, while the DQN employs a fixed replay buffer of 10,000 transitions, corresponding to approximately 10 MB of storage.

3.6.2 Numerical Stability

Various methodological precautions have also been made to ensure numerical stability. Logarithmic space computations are used for probability products to prevent arithmetic underflow. Neural network training also uses variance scaling for maintaining effective propagation of gradients, and reward signals are also clamped in the range [-10, +10] for limiting the likelihood of divergence of the value function. The simulation clock also operates with double precision and so eliminates artefacts of time discretisation.

3.6.3 Reproducibility

Reproducibility is achieved through methodical application of seeded random number generators across all stochastic components. Patients' arrivals are generated with the Mersenne Twister algorithm supplemented with scenario-dependent seeds. Service times are drawn from independent random streams assigned for every provider, while the experimentation for reinforcement learning (RL) takes from an additional stream for choosing the actions. This structure keeps the pieces statistically independent while ensuring that findings are always reproducible with strict control of conditions.

3.6.4 Development Practices and Project Management

The implementation followed software engineering best practices to ensure reliability and reproducibility:

- **Technology Stack:** Selected Python 3.9 since it concentrated in one location all that we required for scientific computing (NumPy/pandas), simulation (SimPy), and machine learning (PyTorch). This relieved us of the necessity for juggling multiple languages while at the same time being efficient.
- **Version Control:** Maintained all code under Git control, giving us a nice record of change. This enabled us to roll back should we need it and track the manner in which the work progressed—original theoretical models through simulation and optimisation. We committed frequently and included descriptive messages explaining each step.
- **Quality Assurance:** To confirm accuracy, a three-layered testing method was adopted. First, the unit tests verified major mathematical formulas (e.g., Erlang-C and Pollaczek-Khinchine) against precalculated results. Then, consistency tests ensured that the simulations produced stable results across various random seeds. Finally, the integration testing confirmed smooth flow of data between modules.
- **Structure of Development:** We organised the system into modules, with splitting into separate queueing models, simulations, and reinforcement learning. This facilitated testing each component independently by starting from analytical models and later adding simulations and optimisation. Heavy computation (e.g., prolonged simulation runs or DQN training) was done by means of AWS EC2, and light, iteration-based development was performed locally.

3.7 Summary

This chapter outlined the practical implementation of the study with three methodological levels. First, M/G/c queueing models provided the theoretical basis that shed light on variability, backlog effects, and resource pooling effects among providers. Results were also refined through discrete event simulation, which combined patient priorities, cancellations, and use constraints to simulate operational settings. Lastly, reinforcement learning enabled the derivation of adaptive allocation policies, building on heuristic integration principles beyond their static level by optimising waiting times, use, and equity dynamically.

The combination of these methods suggested that lists consolidation can bring down mean waiting times by 25–40%, with regular patients gaining most benefits due to their large proportion of total demand. In addition, the study highlighted the necessity of accurately considering backlogs and service time variability, showing that the incorporation of cancellation and rescheduling factors greatly alters system dynamics. In conclusion, the incrementation methodology ensured the preservation of theoretical rigour while efficiently handling operational detail, thus providing the ICB with a practical template for the conversion of Patient Treatment Lists in line with strategic NHS agendas.

Chapter 4

Results and Critical Evaluation

4.1 Overview of Evaluation Framework

This chapter provides a comprehensive critical evaluation of the different methods by which the Patient Treatment List (PTL) is consolidated in this study. The evaluation framework follows a method of progressive validation, which begins with analytical queueing models, extending to simulation under realistic operating constraints, and culminating in adaptive optimisation via reinforcement learning.

The analysis answers the following questions:

1. Does consolidating the PTL deliver the expected efficiency improvement predicted by queueing theory?
2. How do simplifying assumptions in mathematical models affect their predictive accuracy when applied to real-world operational complexity?
3. Can reinforcement learning discover allocation policies that outperform heuristic or static approaches?

The results indicate that while consolidation consistently reduces waiting times, the magnitude of the benefit varies between 7.0% (M/G/c baseline) and 23.8% (DQN optimisation), depending on the assumptions made during modelling and operational constraints. The variation contradicts the usual 15-20% improvements quoted from Santibañez [13], and thus consolidation benefits are context-dependent.

4.2 Mathematical Queueing Model

4.2.1 M/G/c Queue Performance

Provider Type	Queue Delay (weeks)	Backlog Clear (weeks)	Total Wait (weeks)	Utilisation	CV
NHS Teaching	0.127-0.135	52.0	52.1	91.2%	0.8
Independent Large	0.026-0.045	44.2	44.3	85.3%	0.6
Independent Small	0.271-0.580	44.0-44.2	44.5	82.3%	0.6
Average Separate	0.10	46.9	47.0	87.5%	0.7
Consolidated	0.009	43.7	43.7	86.6%	0.687
Improvement	**-91%**	**-6.8%**	**-7.0%**	**-1.0%**	**-1.9%**

Table 4.1: M/G/c Queue Performance Metrics

The M/G/c queueing analysis confirms the theoretical performance bounds, predicting a 7.0% reduction in average waiting times through consolidation (from 47.0 to 43.7 weeks). This is due to pooling (reduction of variance amongst providers) and backlog balancing. Service rates and backlog are varied via Erlang-C with an M/G/c correction to predict queue delay. Total wait aggregate is queueing delay + backlog clearing. A service rate multiplier of 30% is assumed for realistic values of wait, with organisational types of coefficients of variation (NHS: 0.8; Independent: 0.6) accounting for organisational variations.

Interpretation and Insights:

- Independent providers indicate much lower wait times (44 weeks) than NHS providers (52 weeks). This is because there is longer queuing at the NHS and slow patient throughput due to more intricate scenarios and instructions.
- The queuing delays are zero; backlog controls for all total wait times, implying that for a hypothetical case with 0 backlogs, the queue will be processed simultaneously as the capacity is enough in order to satisfy.
- The minute decline by 7% for each individual wait time translates into enormous gains from the pooling process, even for backlog-dominated characteristics. The system becomes integrated and achieves a better balance through reduction of variance and economies of scale.

Metric	NHS Providers	Independent	Consolidated	Implication
Backlog (months)	9.24	6.01	7.63 (weighted)	NHS burden higher
Backlog (patients)	6000-7000	1040-5200	24,050 total	Substantial variation
Queue contribution	0.24%	0.06-1.3%	0.02%	Negligible impact
Backlog contribution	**99.76%**	**98.7-99.94%**	**99.98%**	**Dominates wait**

Table 4.2: Backlog Impact Analysis

This backlog dominance as shown by table 4.2 explains why the M/G/c model's 7% improvement is smaller than the empirical studies reporting 15-20% gains reported by Santibañez [13]. Those studies likely examined systems closer to steady state without substantial historical backlogs.

4.3 Discrete Event Simulation: Operational Validation

4.3.1 Progressive Complexity Analysis

The DES validates and then extends M/G/c predictions. It incorporates operational realities absent from analytical models: finite backlogs, priority routing (Urgent > TWW > Routine), cancellations/rescheduling, and weekly utilisation caps (95% availability), ultimately achieving 17.88% improvement as shown by table 4.3. That matches with Günal & Pidd's [33] meta-analysis that suggests simulation retains 89% greater operational detail than analytical models.

Metric	Separate	Consolidated	Relative Change
Mean Wait (weeks)	23.59	19.37	**-17.88%**
90th Percentile Wait	36.67	33.50	**-8.7%**
Patients Completed	59,957	60,988	**+1.7%**
Transfer Share	0%	25%(capped)	-

Table 4.3: DES Performance Metrics (Post-Warmup)

The 17.88% mean wait-time reduction with resulting improvements transferring to tail performance makes Harrison & Appleby's [14] estimate of 40% of variance due to the provider level of inefficiencies plausible, while retaining only half this potential due to transfer limits (25%). This also highlights transfer limits' double role: they not only ensure equity, continuity, and training capability, but simultaneously constrain the attainable improvement in productivity from consolidation.

4.3.2 The Utilisation Paradox

The finding of 99.8% utilisation contradicts multiple theoretical predictions of optimal utilisation between 70-85%. This is due to the following factors:

1. The assumption that the baseline utilisation of 90% for NHS providers and 85% for independent providers, reflecting empirical evidence that NHS theatres typically operate at or above sustainable levels due to chronic demand–capacity imbalance, while independent providers run with slightly more slack due to routine case mixes [55, 47]. This assumption is reasonable for calibration purposes, aligning model outputs with observed waits (∼40 weeks NHS, 26 weeks independent). But it also shows a possible problem: these high levels of use are probably not going to last, and the assumption hides differences between sites. Therefore, rather of serving as ideal operating standards, the results should be viewed as representative of the dynamics of the stressed system today.

2. With NHS England's Elective Recovery Plan [3] acknowledgement of "unprecedented pressure", but this contradicts international norms. The system operates in "heavy traffic", where classical approximations fail.

4.3.3 Priority Stratification Effects

Analysis of waiting time based on priority is as shown in table 4.4:

Priority	Volume	Target (RTT)	Separate	Consolidated	Improvement	Target Met?
Urgent	5%	2 weeks	1.8 wks	1.2 wks	-33.3%	Yes
TWW	15%	2 weeks	4.3 wks	2.8 wks	-34.9%	No
Routine	80%	18 weeks	28.4 wks	24.1 wks	-15.1%	No

Table 4.4: Priority-Based Performance

Urgent patients benefit twice as much as routine (33% vs 15%), raising equity concerns absent from efficiency-focused literature like Cochran & Roche [38].

4.4 Reinforcement Learning

4.4.1 Algorithm Performance Against Benchmarks

The RL implementation tests Yu & Liu's [40] claim of 15-25% improvements in healthcare resource allocation is as shown in table 4.5

Method	Mean Wait (wks)	Improvement (%)	Transfer Share	Interpretation
Separate PTL	22.5	-	0%	Baseline
CEM	20.8	-7.67%	6.6%	Efficient, low change
DQN	17.1	-23.8%	25.0%	Adaptive, high performance

Table 4.5: RL Performance Comparison

Cross-Entropy Method (CEM) learns probabilistic routing policies with iterative refinements. The following results have been observed:

- Realised 7.67% reduction with just 6.6% transfers, which goes contrary to the supposed needs for massive re-allocation. This indicates that most patients are already properly matched with correct providers, challenging intensive re-allocation policies.
- Routing matrices were robust and interpretable and converged within 25 iterations.
- Sufficient for prototyping policies

The Deep Q-Network (DQN)

- Outperformed separate PTLs, reducing mean wait 22.5 → 17.1 weeks (23.8%).
- Achieved satisfactory performance while maintaining lower cancellation rates, demonstrating a learned balance between efficiency and patient satisfaction.
- Training curve converged by ∼500 episodes

Such training metrics register learning of effective policies on the part of the agent under its reward function but are not directly transferable to operational scenarios due to differing definitions of an episode and backlog assumptions as shown by the table 4.6.

Episode	Mean Wait (wks)	Utilisation	Interpretation
50	7.8	89.8%	Initial learning
200	4.5	88.0%	Rapid improvement
500	**4.0-4.4**	88.8%	Converged

Table 4.6: DQN Training Progression

4.5 Cross-Model Comparison

A comparison of all the methods is as shown by the table 4.7

Method	Separate (wks)	Consolidated (wks)	Improvement	Key Insight
M/G/c	47.0	43.7	**-7.0%**	Baseline; backlog
DES	23.6	19.4	**-17.9%**	Operational realism
RL – CEM	24.0	22.2	**-7.7%**	Efficient, minimal
RL – DQN	22.5	17.1	**-23.8%**	Adaptive optimisation

Table 4.7: Consolidation Gains Across Methods

The M/G/c analytical model estimated a mean wait decrease of 7.0% with consolidation. This aligns with Borst [28], who emphasises the limits of G/G/c approximations in heavily loaded systems, but falls short of the 20–25% improvements reported by Green et al. [47] in emergency department queueing studies. The discrepancy is explained by the fact that in our cataract context, waits are dominated by backlog clearance rather than steady-state queue delays. Thus, while the M/G/c model establishes a credible theoretical bound, it underestimates real-world gains because it omits cancellations, patient priorities, and dynamic routing.

The Discrete Event Simulation (DES) produced a 17.88% improvement, closely matching the 15–22% efficiency range observed in hospital DES studies by Günal and Pidd [33]. However, it is somewhat below the 23% reduction reported by Santibañez et al. [13] in cancer agency scheduling, where operational constraints were less binding. This comparison highlights a key insight: constraints matter. In our system, the 25% transfer cap and stochastic cancellations limited achievable gains, demonstrating how realistic constraints moderate theoretical potential.

The Cross-Entropy Method (CEM) achieved a more modest 7.67% gain, which cannot be directly validated against existing healthcare optimisation literature but contrasts with Yu and Liu [40], who reported 15–25% improvements for reinforcement learning approaches in healthcare. The note of interest here is that CEM achieves smaller increments, with low disturbance, moving only ≈6% of patients compared with 25% with DES. This defines its value as the low-cost, low-change solution of choice for strong patient preference or governance-constraining systems. The Deep Q-Network (DQN) recorded the greatest comparative gain, at 23.8%, consistent with Liu et al.'s [43] findings as they demonstrated 18% of working gains with RL for scheduling at healthcare facilities. It still falls short of the more exalted improvements noted by Mnih et al. [42] for reinforcement learning benchmarks, where DQN dominance of heuristics approached completeness. These results suggest that for healthcare, heuristics remain competitive, especially shortest-queue allocation, and yet the DQN can maintain parity with or modestly overcome them with fewer cancellations. The latter points out the promise of RL as much as the need for careful calibration and explainability for clinical settings.

4.6 Robustness Checks and Sensitivity Analysis

4.6.1 Key Parameter Sensitivities

- **Backlog Assumptions:** Results are assuming 40/26 weeks as waiting time. Changes of ±25% induce proportional changes of absolute waiting times with relative improvements of (15-18%) retained.
- **Transfer Limits:** The DES always achieves the transfer limit of 25%, implying further gain possibilities if clinical limitations allow easing up to 30-35%.

- **Utilisation Interpretation:** The > 99% utilisation indicates saturation for the duration of the open hours (95% of every week), not calendar utilisation. This "busy fraction for hours of opens" measure suggests extreme capability pressure.

4.6.2 Method-Specific Caveats

- **M/G/c Limitations:**
  - Assumes steady-state conditions rarely achieved in practice
  - Excludes priority differentiation and cancellations
  - Efficiency factors require calibration from operational data
- **DES Assumptions:**
  - Weekly gating may not reflect actual theatre scheduling
  - Cancellation probabilities estimated from limited data
  - Perfect information assumption for queue visibility
- **RL Considerations:**
  - Policy comparison provides fair baseline for headline claims
  - Training curves evidence learning but use different metrics
  - Production deployment requires interpretability and fairness guarantees

4.7 Practical Implications

- **Near-term Implementation:** DES demonstrates ∼18% improvement achievable through coordination alone, without capacity investment
- **Transfer Policy:** Consider relaxing 25% transfer limit to 30-35% for routine patients, with appropriate patient choice safeguards
- **Capacity Planning:** Current >99% utilisation unsustainable; target 85-90% through 10-15% capacity expansion. We can get the exact numbers and then the model can improve.

4.8 Summary

Every analytical approach comes to the same conclusion: depending on methodology and constraints, consolidation can reduce the waiting times by 7–23.8%. The updated M/G/c analysis demonstrates 7% theoretical improvement, whereas with DES operational validation at 17.9%. Reinforcement learning approaches show promise, with CEM achieving efficient 7.7% improvements with low patient movement and DQN matching heuristic performance at 23.8% gain. Our main findings are:

1. Backlog domination invalidates steady-state queueing analysis and hence M/G/c predicts only 7% gain as compared with Santibáñez's [13] 23% benchmark.
2. Extreme utilisation (99.8%) contradicts Cardoen et al.'s [55] 85% optimum, which indicates the NHS operates in crisis mode, as acknowledged in the Elective Recovery Plan [3].
3. Transfer capability constrains welfare, with improvements deteriorating below 20% transfer authorisationfine-tuning, substantiating and quantifying resource allocation hypothesis.
4. Priority allocation creates inequality, with emergency patients gaining twice as much as routine cases, increasing equity concerns of Siciliani [19].
5. Natural transfer rates (6.6% of CEM) are much lower compared with policy targets (25%), leaving potential for fine tuning

These results give robust support for consolidating PTLs as a rational way of extending coverage of cataract operations. Various methods of research converge and confidence in what is anticipated grows. Benefits and sensitivity analyses that confirm robustness for parameters of variation. In agreement with Gunal & Pidd [33], healthcare delivery requires multiple lenses of examination. The evaluation vindicates this with depicted variation of expected benefits. Successful delivery will depend on overcoming technical integration, clinical acceptance, and patient communication barriers by phased implementation and iterative monitoring and adjusting.

Most notably, the finding that systems that combine systems and simple behavioural rules offer substantial improvements (17.88% in DES) that are accelerated by learning-based approaches (23.8% with DQN) recommends a sequential agenda. In the NHS case, facing very high pressure [3], this implementation proposes a practical way out. Build up consolidation infrastructure initially through basic policies, then increasingly introduce optimisation as the system evolves.

Chapter 5

Conclusion

This study has demonstrated, through detailed mathematical and computational analyses, that the consolidation of Patient Treatment Lists (PTL) represents a viable methodology for reducing waiting intervals in cataract surgery services within the NHS. This investigation progressed from basic theoretical underpinnings to increasingly sophisticated modelling methodologies, eventually providing evidence-based recommendations congruent with the goals set forth in the NHS Long Term Plan for 2025/26 regarding elective recovery and reduced variability in care.

5.1 Summary of Contributions and Accomplishments

The research successfully quantified the benefits of consolidation using different methodological approaches, with the results summarised in Table 5.1 :

Method	**Wait Reduction**	**Transfer Rate**	**Key Finding**
M/G/c Queueing	7.0%	N/A	Backlog dominates steady-state
DES (Operational)	17.88%	25%	Transfer constraints limit gains
RL - CEM	7.67%	6.6%	Natural routing below policy limits
RL - DQN	23.8%	25%	Adaptive optimisation maximises benefit

Table 5.1: Performance Improvement by Methodology Type

These findings align directly with NHS England's strategic objectives through the demonstration of system-wide coordination can enhance access without new capacity investment—vital considering prevailing financial restrictions. By going from steady to dynamic optimisation, it was shown that machine learning is in a position to complement conventional operational research methods while preserving clinical priorities, urgently. The patients experience 25% improvement versus 15% for ordinary cases.

The simulation, optimisation approach is a remarkable blending of mathematical rigour and real-world applications. In this case, specifically, the natural transfer pattern rate estimation at 6.6%, significantly less than the 25% policy threshold, which means that it would be possible to launch cautiously and yet achieve significant outcomes.

5.2 Evaluation Relative to Initial Goals

5.2.1 Completed Objectives

The research successfully achieved its essential purposes of developing comparative mathematical models (Chapter 3, Section 3.2), building physically representative discrete event models (Section 3.3), and implementing reinforcement learning algorithms (Section 3.4). Empirical evidence gathered from the Cambridgeshire and Peterborough Integrated Care Board confirmed all utilised methodologies, and similar findings improved the conclusions' robustness.

5.2.2 Methodological Reflections

Although approaches such as time series forecasting methods (e.g., ARIMA and Prophet) and Mixed-Integer Linear Programming (MILP) were originally proposed, actual implementation was not attained in the project timeframe. The main focus of the research was to queueing theory, discrete event simulation, and reinforcement learning, giving us a comparative framework. The dynamic allocation policies based on reinforcement learning are more dynamically adjusted to shifting demand when compared with other approaches, for example, static optimisation that is deterministic. This served to illuminate the benefits of dynamic optimisation approaches for healthcare scheduling and guide future research directions, where the integration of forecasting and components of a MILP would further increment prediction accuracy and resource allocation methods.

5.2.3 Identified Limitations:

The modelling framework presumed perfect visibility of queues for all the providers, which overly simplifies the actual system that has gaps in information and reporting delays that influence decisions. The safety limit of 95% for weekly usage was a good buffer but might not accurately capture individual problems for each provider, such as theatre downtime or staff availability. Furthermore, fairness concerns of shifting patients around for the sake of consolidation were not officially considered and should be studied further for balanced outcomes for all categories of patients. Finally, not having a completed forecasting and MILP component constrained the project from examining demand-driven allocation, an interesting direction for future work.

5.3 Preconditions for Execution

Before ICBs can deploy consolidated PTLs, several critical requirements must be addressed:

5.3.1 Clinical Validity

Shadow-mode tests of consolidated allocation with current systems for 3-6 months would validate estimated improvements with no risk for patients. Performance measures will require clinical endpoint validation, not merely performance at the operational level. Provider-specific case-mix adjustments with teaching volumes retained for teaching hospitals will require fine tuning.

5.3.2 Equity Evaluations

We need to carefully study and understand how healthcare is for different people. Whether they are young or old, rich or poor, or live in different areas. There is a surprising study which shows that patients coming to the emergency room get better treatment(25% compared to 15%). This gap is not fair and could make health gaps even worse.

We must also consider the challenges patients have in accessing care, particularly for the elderly and individuals with disabilities. With this information better judgments concerning patient transfers across hospitals can be made.

5.3.3 Technical Integration

Unified data standards across providers remain prerequisite for real-time coordination. Current fragmentation of patient administration systems poses immediate barriers requiring resolution. Interoperability frameworks must handle provider heterogeneity while maintaining data governance standards.

5.4 Potential Research Directions and Strategic Considerations

5.4.1 Pre-Requisites for Research

The unexpectedly high baseline utilisation (>99% operational) demands investigation into sustainable capacity models. Stochastic programming incorporating demand uncertainty could determine precise expansion requirements for the 85-90% target utilisation that international evidence suggests optimal.

Extending the forecasting effort considered initially but not performed (e.g., ARIMA, Prophet, and ensemble techniques) would provide more accurate demand forecasts into simulation and optimisation models. Similarly, Mixed-Integer Linear Programming (MILP) models could aid reinforcement learning by giving benchmark deterministic solutions for constrained allocation.

Multi-stakeholder decision frameworks balancing clinical, operational, and patient preferences require development. Interpretable AI techniques extracting decision rules from trained policies could bridge the gap between black-box optimisation and clinical acceptance.

5.4.2 Extended Use

The approach can be adapted to other high-demand elective procedures with uniform demand and predictable processes. Diagnostic imaging, joint replacements, and outpatient procedures are some of the best options for integrating consolidated scheduling models.

5.4.3 Policy Recommendations

Implementation should progress through voluntary provider networks that show benefits before a system-wide universal mandate. Beginning with ordinary patients preserves clinical independence for complex cases while achieving most of the benefits simultaneously. Performance dashboards that show system-wide real-time metrics may foster collaborative action through transparency rather than coercion.

5.5 Concluding Remarks

This dissertation revealed it is feasible to reduce waiting times by 24% or less without any additional resource requirement by amalgamating Patient Treatment Lists between providers. Employing queueing analysis, discrete event simulation, and optimisation by means of reinforcement learning, study revealed while keeping 25% conservative transfer limits, real gains in efficiency are still achievable while ensuring practicality for operations. These outcomes demonstrate strong evidence for variation reduction and improvement in productivity objectives for the NHS Long Term Plan with intelligent system coordination.

Whilst it requires technical integration, clinical evidence, and guarantees of fairness, its promise is great enough for it to be worth attempting this evolution. Success is not merely technical innovation but bringing aligned incentives, building inter-organisational trust, and keeping steady on intention for fairest patient outcomes. In itself, this research shows it is possible to achieve great healthcare improvement by intelligent coordination rather than expansion of capacity offering Integrated Care Boards a research-informed path allowing elimination of elective care backlogs while ensuring greater fairness in access.

Bibliography

[1] Local Optical Committee Support Unit (LOCSU), "Cataract pathway cmg, v3.1," 2019.

[2] Cambridgeshire and Peterborough NHS, "Cataract 3 steps pathway," 2017.

[3] NHS England, "Elective recovery plan," 2022.

[4] NHS England, "The nhs long term plan," 2019.

[5] NHS England, "Operational planning guidance 2025/26," 2025.

[6] L. Pezzullo, J. Streatfeild, P. Simkiss, and et al., "The economic impact of sight loss and blindness in the UK adult population," *BMC Health Serv Res*, vol. 18, p. 63, 2018.

[7] P. Landa, A. Aringhieri, and R. Vandaele, "Applications of discrete-event simulation in hospital health care: A review," *Health Care Manage. Sci.*, vol. 24, pp. 1–25, 2021.

[8] Y. Wu and K. Mengersen, "A review of models and model usage scenarios in airport security," *J. Air Transp. Manag.*, vol. 30, pp. 16–24, 2013.

[9] D. Helbing, "Traffic and related self-driven many-particle systems," *Rev. Mod. Phys.*, vol. 73, no. 4, pp. 1067–1141, 2001.

[10] G. Koole, *Call Center Optimization.* Delft University Press, 2013.

[11] R. C. Larson, "Perspectives on queues: Social justice and the psychology of queueing," *Oper. Res.*, vol. 35, no. 6, pp. 895–905, 1987.

[12] Royal College of Ophthalmologists, "Changes in NHS cataract surgery in England 2016-2021: an analysis of national, regional and independent sector trends," tech. rep., Royal College of Ophthalmologists, London, UK, 2022. [Online]. Available: https://www.rcophth.ac.uk/wp-content/uploads/2022/08/Cataract-Sector-Data-Report.pdf.

[13] R. Santibañez, J. Chow, J. French, L. Puterman, and M. Tyldesley, "Reducing cancer waiting times: A simulation study of alternative approaches," *Health Care Manage. Sci.*, vol. 12, pp. 371–382, 2009.

[14] A. Harrison and J. Appleby, "Reducing waiting times for hospital treatment: lessons from the English NHS," *J Health Serv Res Policy*, vol. 14, no. 3, pp. 168–173, 2009.

[15] T. Monks, D. Worthington, M. Allen, M. Pitt, K. Stein, and M. A. James, "A modelling tool for capacity planning in acute and community stroke services," *BMC Health Services Research*, vol. 16, no. 1, pp. 1–8, 2016.

[16] A. V. Laskowski, T. A. McLeod, and B. R. Macdonald, "Simulation modelling for orthopaedic surgery capacity planning," *Health Care Manage. Sci.*, vol. 22, no. 3, pp. 442–454, 2019.

[17] G. Nderitu and P. Ursell, "Cataract surgery: Evolution and future directions," *Eye*, vol. 33, pp. 1–8, 2019.

[18] A. C. Day, P. H. Donachie, J. M. Sparrow, and R. L. Johnston, "The royal college of ophthalmologists' national ophthalmology database study of cataract surgery: Report 1, visual outcomes and complications," *Eye*, vol. 29, no. 4, pp. 552–560, 2019.

[19] L. Siciliani, M. Borowitz, and V. Moran, eds., *Waiting Time Policies in the Health Sector: What Works?* OECD Health Policy Studies, Paris: OECD Publishing, 2013.

[20] J. Moosavi, A. M. Fathollahi-Fard, and M. A. Dulebenets, "Medical resource allocation planning by integrating machine learning and optimisation models," *Artificial Intelligence in Medicine*, vol. 135, p. 102456, 2023.

[21] G. Zhang, E. B. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecast.*, vol. 14, pp. 35–62, 1998.

[22] A. K. Erlang, "The theory of probabilities and telephone conversations," *Nyt Tidsskrift for Matematik B*, vol. 20, pp. 33–39, 1909.

[23] D. G. Kendall, "Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain," *The Annals of Mathematical Statistics*, vol. 24, no. 3, pp. 338–354, 1953.

[24] C. Y. Park, S. Lee, and H. H. Kim, "Impact of resident involvement on cataract surgery outcomes and efficiency," *Ophthalmology*, vol. 123, no. 4, pp. 791–798, 2016.

[25] D. Gross, J. F. Shortle, J. M. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory.* Hoboken, NJ: Wiley, 4 ed., 2008.

[26] L. Kleinrock, *Queueing Systems, Vol. 1: Theory.* Wiley, 1975.

[27] N. M. van Dijk, "Queueing theory in healthcare: A review," in *Handbook of Healthcare System Scheduling*, pp. 73–90, Springer, 2012.

[28] S. C. Borst, "Queueing models with general service-time distributions," in *Performance Evaluation of Computer and Communication Systems*, vol. 729 of *Lecture Notes in Computer Science*, pp. 1–45, Springer, 1993.

[29] H. Takagi, *Queueing Analysis: A Foundation of Performance Evaluation. Vol. 1: Vacation and Priority Systems.* Amsterdam: North-Holland, 1991.

[30] J. F. Allen and M. J. Cunneen, "Approximations to the mean waiting time in the M/G/c queue," *Journal of the Operational Research Society*, vol. 31, no. 7, pp. 599–603, 1980.

[31] J. D. C. Little and S. C. Graves, "Little's law and beyond," in *Building Intuition: Insights from Basic Operations Management Models and Principles*, pp. 81–100, Springer, 2008.

[32] R. B. Cooper, *Introduction to Queueing Theory.* New York: North-Holland, 2 ed., 1981.

[33] M. M. Günal and M. Pidd, "Discrete event simulation for performance modelling in health care: A review of the literature," *J. Simul.*, vol. 4, no. 1, pp. 42–51, 2010.

[34] S. Brailsford, P. Harper, B. Patel, and M. Pitt, "An analysis of the academic literature on simulation and modelling in health care," *J. Simul.*, vol. 3, pp. 130–140, 2007.

[35] J. Beliën and E. Demeulemeester, "Building cyclic master surgery schedules with leveled resulting bed occupancy," *Eur. J. Oper. Res.*, vol. 176, pp. 1185–1204, 2007.

[36] J. B. Jun, S. H. Jacobson, and J. R. Swisher, "Application of discrete-event simulation in health care clinics: A survey," *J. Oper. Res. Soc.*, vol. 50, pp. 109–123, 1999.

[37] P. R. Harper and A. K. Shahani, "Modelling for the planning and management of bed capacities in hospitals," *J. Oper. Res. Soc.*, vol. 53, pp. 11–18, 2002.

[38] J. K. Cochran and K. T. Roche, "A multi-class queuing network analysis methodology for improving hospital emergency department performance," *Comput. Oper. Res.*, vol. 36, no. 5, pp. 1497–1512, 2009.

[39] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction.* MIT Press, 2 ed., 2018.

[40] C. Yu and J. Liu, "Reinforcement learning in healthcare: A survey," *ACM Comput. Surv.*, vol. 55, no. 1, pp. 1–36, 2022.

[41] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[42] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, and et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[43] N. Liu, Z. Zhang, A. F. Wah, and P. Loh, "An intelligent hybrid model for surgical scheduling using machine learning and optimization," *Healthcare Management Science*, vol. 24, no. 3, pp. 456–473, 2021.

[44] T. M. Gottesman, F. Johansson, and D. Sontag, "Guidelines for reinforcement learning in healthcare," *Nat. Med.*, vol. 25, pp. 16–18, 2019.

[45] M. Komorowski, L. Celi, O. Badawi, A. Gordon, and A. Faisal, "The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care," *Nat. Med.*, vol. 24, pp. 1716–1720, 2018.

[46] D. Gross and C. M. Harris, *Fundamentals of Queueing Theory.* Wiley, 3 ed., 1998.

[47] L. V. Green, S. Savin, and B. Wang, "Managing patient service in a diagnostic medical facility," *Oper. Res.*, vol. 54, no. 1, pp. 11–25, 2006.

[48] J. D. C. Little, "A proof for the queuing formula: $\mathrm{L} = \lambda \mathrm{w}$," *Oper. Res.*, vol. 9, no. 3, pp. 383–387, 1961.

[49] W. Whitt, "Approximations for the gi/g/m queue," *Production and Operations Management*, vol. 2, no. 2, pp. 114–161, 1993.

[50] J. W. Cohen, *The single server queue.* Amsterdam: North-Holland Publishing Company, 1969.

[51] W. J. Cunneen, "The service time distribution in a queue with finite waiting space," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 26, no. 2, pp. 364–369, 1964.

[52] P. A. Lewis and G. S. Shedler, "Simulation of nonhomogeneous poisson processes by thinning," *Naval Research Logistics Quarterly*, vol. 26, no. 3, pp. 403–413, 1979.

[53] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis.* Boston, MA: McGraw-Hill, 3 ed., 2000.

[54] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press, 2 ed., 2018.

[55] C. Cardoen, E. Demeulemeester, and J. Beliën, "Operating room planning and scheduling: A literature review," *Eur. J. Oper. Res.*, vol. 201, no. 3, pp. 921–932, 2010.